Haptify: A Measurement-Based Benchmarking System for Grounded Force-Feedback Devices

Farimah Fazlollahi, *Graduate Student Member, IEEE*, and Katherine J. Kuchenbecker, *Fellow, IEEE*

***Abstract*—Grounded force-feedback (GFF) devices are an established and diverse class of haptic technology based on robotic arms. However, the number of designs and how they are specified make comparing devices difficult. We thus present Haptify, a benchmarking system that can thoroughly, fairly, and noninvasively evaluate GFF haptic devices. The user holds the instrumented device end-effector and moves it through a series of passive and active experiments. Haptify records the interaction between the hand, device, and ground with a seven-camera optical motion-capture system, a 60-cm-square custom force plate, and a customized sensing end-effector. We demonstrate six key ways to assess GFF device performance: workspace shape, global free-space forces, global free-space vibrations, local dynamic forces and torques, frictionless surface rendering, and stiffness rendering. We then use Haptify to benchmark two commercial haptic devices. With a smaller workspace than the 3D Systems Touch, the more expensive Touch X outputs smaller free-space forces and vibrations, smaller and more predictable dynamic forces and torques, and higher-quality renderings of a frictionless surface and high stiffness.**

***Index Terms*—Benchmarking, grounded force-feedback devices, haptics, performance metrics.**

I. Introduction

OVER the last 30 years, researchers have invented hundreds of haptic interfaces that enable a human user to touch a virtual or remote environment. Common applications include education, computer-aided design, gaming, surgery, and hazardous material handling. These devices vary widely in their mechanical design, what they sense and actuate, how they are controlled, and other factors that consider both technical requirements and human insights. As the first main type of haptic interface to be developed, a *grounded force-feedback (GFF) device* is a mechatronic system that is mounted to a stationary surface such as a table and that measures the user's motion and/or force, and outputs forces and/or motions in response [1], [2]. GFF devices vary from simple haptic paddles and knobs to high-degree-of-freedom robotic arms (see Fig. 1).

According to Massie and Salisbury's [3] pioneering 1994 paper on the first successful GFF device (the PHANToM), the *three necessary criteria for an effective haptic interface* are: 1) free space must feel free, 2) solid virtual objects must feel stiff, and 3) virtual constraints must not be easily saturated. Designers thus seek to create mechanisms that are easily backdrivable, stiff, and strong. Haptipedia is an open-source taxonomy and database that has recently visualized the specifications of more than 100 GFF devices from the literature and commercial product datasheets [4].

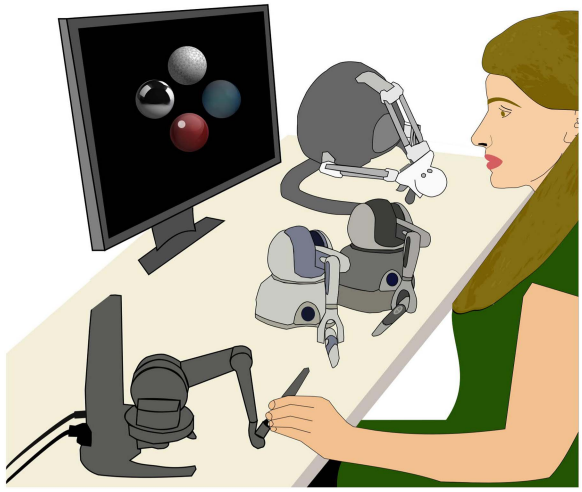

Fig. 1. User holds the end-effector of a GFF device to touch four virtual spheres. The forces she feels are greatly affected by not only the rendering capability of the device but also the device's own dynamics, such as its workspace edges, weight, and friction.

Although many GFF haptic devices exist, there is no standard framework for reporting their performance, and many reported metrics are not well-grounded in user opinion. In a recent study Seifi et al. [5] showed that common definitions of device attributes do not correspond well with device affordances; expert opinions about device capabilities go far beyond the low-level specifications currently reported in papers, data sheets, and Haptipedia. Instead, experts typically rely on their experience physically testing devices.

Most other researchers concerned with haptic interface quality have focused on individual devices. For example, the kinematics, manipulability, and gravity compensation of three devices from the same company (SensAble, now 3D systems) were separately characterized by different research teams: the Phantom Omni (now Touch) [6], [7], the Phantom Desktop (now Touch X) [8], and the Phantom Premium [9]. Samur [10] and Hatzfeld and Kern [11] have both made progress on defining and calculating performance metrics for a few haptic interfaces. However, these

Manuscript received 26 March 2022; revised 29 September 2022; accepted 9 November 2022. Date of publication 16 December 2022; date of current version 5 April 2023. This research was supported by the Max Planck Society. This paper was recommended for publication by Associate Editor J. Ueda and Editor P. Robuffo Giordano upon evaluation of the reviewers' comments. *(Corresponding author: Farimah Fazlollahi.)*

The authors are with the Haptic Intelligence Department, Max Planck Institute for Intelligent Systems, 70569 Stuttgart, Germany (e-mail: fazlollahi@is.mpg.de; kjk@is.mpg.de).

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TRO.2022.3226110.

Digital Object Identifier 10.1109/TRO.2022.3226110

metrics almost all depend on the device's own sensor readings, which may not be accurate and certainly do not capture everything a user can feel. Indeed, as Samur [10] presciently stated, the most difficult challenge in characterizing devices is not in defining good metrics; instead, it is in practice, since the applicability of measurements and calculation of meaningful results are highly limited by the lack of proper experimental equipment, nonlinear device behavior, and position dependency.

As more robotic technologies are developed, scientific methodologies for standardization and benchmarking are essential to speed up and strengthen their wide acceptance [12]. New benchmarks have been developed for several subfields of robotics as they mature and expand. In the field of human–robot interaction, Aly et al. [13] presented a subjective analysis of metrics, benchmarks, frameworks, and recent technologies in cognitive robotics, and Feil-Seifer et al. [14] presented the existing benchmarks for evaluating socially assistive robotics. RoboBench [15] is a platform for sharing simulations of full robotic systems, aiming to make a sustainable benchmark for reproducible research. Standard benchmarks are also crucial to the improvement of work in other interdisciplinary fields, as they create reproducible measurements that can quantify progress [15].

There is currently no standard benchmarking system that can evaluate many GFF devices and compare them from the user's perspective. Understanding how each device feels when one performs different tasks, e.g., slow motions in free space, fast oscillations, or exploring a particular rendered virtual environment, requires a standard method for recording and modeling the haptic interaction in the same way for all devices. Only a few previous studies have used haptic sensors to compare how real objects and rendered objects feel to the user. For example, Culbertson and Kuchenbecker [16] investigated the realism of data-driven virtual surfaces rendered with a Phantom Omni. Their force and acceleration measurements showed that the device's passive mechanical properties negatively affected the rendered friction forces, tapping transients, and texture vibrations; the user feels a distorted version of the commanded haptic signals both in free space and when touching virtual objects [16]. The mechanical properties of the hardware also affect what the user feels in teleoperation, where the goal is haptic transparency. Kuchenbecker and Niemeyer [17], [18] considered how mechanism friction and compliant connecting elements affect a GFF device's dynamic behavior, presenting a painstaking and invasive method for developing an accurate, high-order, nonlinear model of these dynamics. Most relevant to our work, Frisson et al. [19] proposed to create an open-source platform for benchmarking haptic devices leveraging the robot operating system (ROS). Their current prototype of RepHap measures interaction forces while using a two-DOF gantry to move the device's end-effector in the left-right and up–down directions and using a third rail to move the haptic device base inward and outward. This approach is limited by the gantry's workspace size, lacks any rotational DOF, and will measure the gantry's own vibrations. Moving the device base also causes the force readings to include the inertial forces caused by the base's mass.

Here, we present and validate *Haptify* (see Fig. 2), a new comprehensive benchmarking system for GFF devices. Haptify uses external sensors to measure the performance of a GFF device while a user holds the device end-effector and performs motions. We define new metrics for evaluating GFF device performance based on external sensor measurements from interaction experiments, and we compare how two devices feel to the user during similar tasks. Our approach is inspired by real use cases of GFF devices, in which the device is mounted on a table, here a force plate, and the human user holds the device end-effector and can move throughout the device workspace, in all the ways allowed by the device. Section II explains the essential design criteria that drove the creation of Haptify. Section III characterizes the three main parts of the setup: the motion-capture system, the force plate, and the sensing end-effector. Section IV introduces the six benchmarking methods that we use to analyze a device's passive and active performance. Section V explains how we collected data from two commercial GFF devices, a Touch and a Touch X. Section VI presents our experimental results; a narrated video about Haptify is also available in the supplementary materials. Section VII discusses the results, and Section VIII explains the implications of our findings, the limitations of our approach, and our plans for future research.

II. System Design

To design Haptify, we went through the commercial devices listed on Haptipedia [4] to determine the type and working range of the sensors that would be required. Like robotic arms, GFF devices can be approximately modeled as multibody dynamic systems; see [20] for a good review. Each link has mass, rotational inertia, and weight due to gravity. Joints transmit both passive and active forces and torques between links and also incur friction and vibration. The user applies forces and torques at the end-effector, and the ground applies forces and torques to keep the base stationary. Therefore, separate sensors are required to know the interaction at the end-effector (between the user and the device) and the ground (between the device and the base), as well as to know the dynamic states of each link throughout the experiments.

Haptify aims to be minimally invasive to avoid damaging the device or altering how it feels. Thus, an accurate optical motion-capture system with small passive markers is a good choice for recording the positions and velocities of device links. Except for the Haption Inca, which is a room-sized cable-driven haptic device, the workspace size of all commercial haptic devices in Haptipedia are in the range of 1.5 m$\times$1.5 m$\times$1.5 m. Therefore, this is the minimum required working volume for our motion-capture system. A Vicon system with seven Vantage V5 infrared cameras and one Vue HD video camera was chosen as it can cover the required volume with 1 mm accuracy and a sampling rate of 100 Hz. Its 3 mm hemispherical markers can adhere to the links and have negligible mass (0.02 g).

For measuring the forces and torques at the base, the required range comes from the Haption Virtuoso 6-D device with a weight of 120 N. It can output a maximum force of 70 N and maximum torque of 5 N $\cdot$ m, which are the largest among the commercial

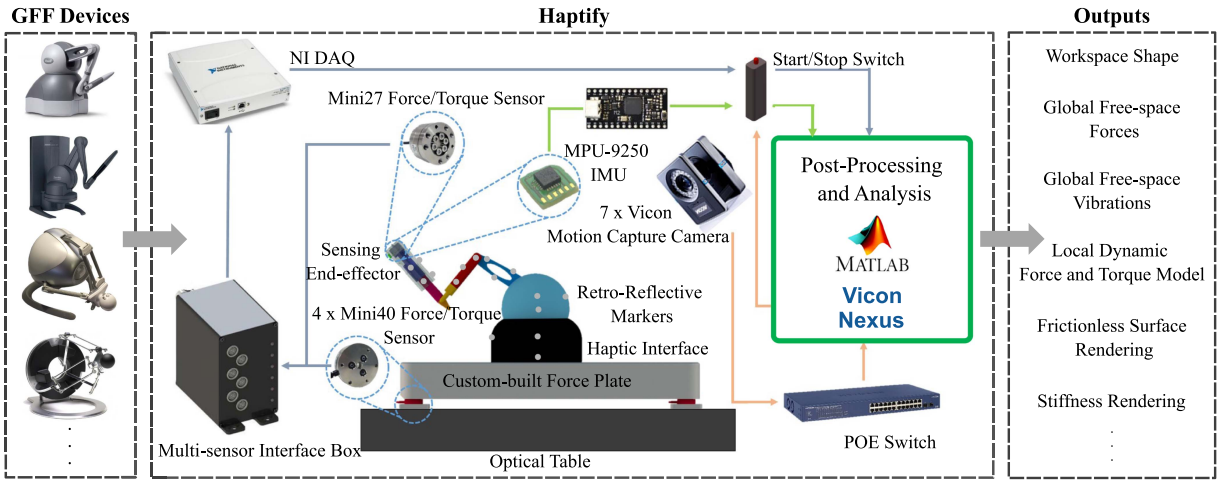

Fig. 2. The Design of Haptify. The GFF device being examined is brought to Haptify and placed on the force plate. The sensing end-effector is attached to the device's handle, and retro-reflective markers are attached to each of its links. The examiner holds the end-effector and moves it in different ways; Haptify records the interaction forces, torques, and vibrations at the handle and the base, as well as the movement of each link. The measurements recorded when the device is unpowered are processed to output the shape of the device's workspace, global free-space forces, global free-space vibrations, and a local dynamic force and torque model. Haptify also measures frictionless surface rendering and stiffness rendering quality when the device is active.

devices. Since the directions of the peak force and torque are not mentioned in the data sheet, we choose a required force range of 70 N in the horizontal plane and 190 N along the vertical axis. The required torque range is 5 N · m around all axes. The required size for the sensor at the ground comes from the Force Dimension Sigma.7 device, which has the largest base size of about 50 cm square among the commercial devices. To fulfill these requirements, we built a rectangular 60 cm ×60 cm force plate with four ATI Mini40 (SI-40-2) force/torque sensors at the corners. The working range of the force plate is ±160 N in the horizontal plane and ±480 N in the vertical direction; the bias weight of the plate (160 N) yields a total vertical measuring range from –320 N to +640 N. Given its wide base, it can measure more than ±100 N· m torque around its center point in all directions.

The required force and torque range for the sensor at the end-effector is the same as the ground sensor without the full device weight. However, since this sensor needs to be attached to the device end-effector, the sensor shape, form factor, and weight are also important criteria. An ATI Mini27 Titanium (SI-80-4) force/torque sensor was chosen for the end-effector sensor. Its working range is ±80 N in the lateral directions, and ±160 N in the axial direction; it can sense a maximum torque of ±4 N · m about the lateral directions and ±2 N · m in the axial direction. The resonant frequency of this sensor is 4900 Hz, which is far higher than human movement bandwidth and also higher than human haptic sensing capabilities. Because parasitic vibrations felt by the user strongly influence the haptic interaction quality [16], and because handle acceleration cannot be estimated accurately from motion-capture measurements, a digital TDK InvenSense MPU-9250 accelerometer with a working range of 40 m/s^2 is also included on the end-effector.

A National Instruments USB-6255 DAQ reads all the analog voltages from the ATI multisensor interface box (9105-IFPSMC-6) that collects inputs from the five force/torque sensors. The accelerometer data are read through a custom-built M2 microcontroller built around the high-performance ATmega32U4 processor. We synchronize all data lines using a multicontact switch to send a trigger signal to the motion-capture system, DAQ, and microcontroller. The motion-capture data are recorded through the Vicon Nexus software, and the DAQ and microcontroller data are recorded via MATLAB; the recordings are all processed offline in MATLAB.

TABLE I
REQUIRED SENSOR TYPES, THEIR RESPECTIVE WORKING RANGES, AND IMPORTANT DESIGN CRITERIA

Sensor		Position	Force	Torque	Vibration
Range	horizontal	±1.5 m	±70 N	±5 Nm	±40 m/s^2
	vertical	±1.5 m	±190 N	±5 Nm	±40 m/s^2
Criteria	data quality (precision, resolution, accuracy, bandwidth), low weight, adaptability to devices, compatibility with the environment, form factor, and availability.				

Table I summarizes all the sensor requirements suitable for fully benchmarking 27 of the 30 (90%) commercial devices on Haptipedia. Beyond the room-sized Haption Inca, the remaining two devices have higher force or torque output, like the Moog FCS HapticMaster [21], which can output a translational peak force of 250 N. Such devices can be still examined by Haptify while moving passively and actively outputting lower forces and torques. Haptify can also be used to benchmark the vast majority of custom GFF devices listed in Haptipedia. Though capable, the selected sensors are admittedly costly (approximately 175 000 USD total); we envision this first version of Haptify as a unique precision instrument that can gather detailed benchmarking data from a wide range of devices while also allowing us to determine whether a simpler set of sensors could yield similar insights.

III. System Characterization

This section characterizes the three main sensing components of Haptify, which are the motion-capture system, the force plate, and the sensing end-effector.

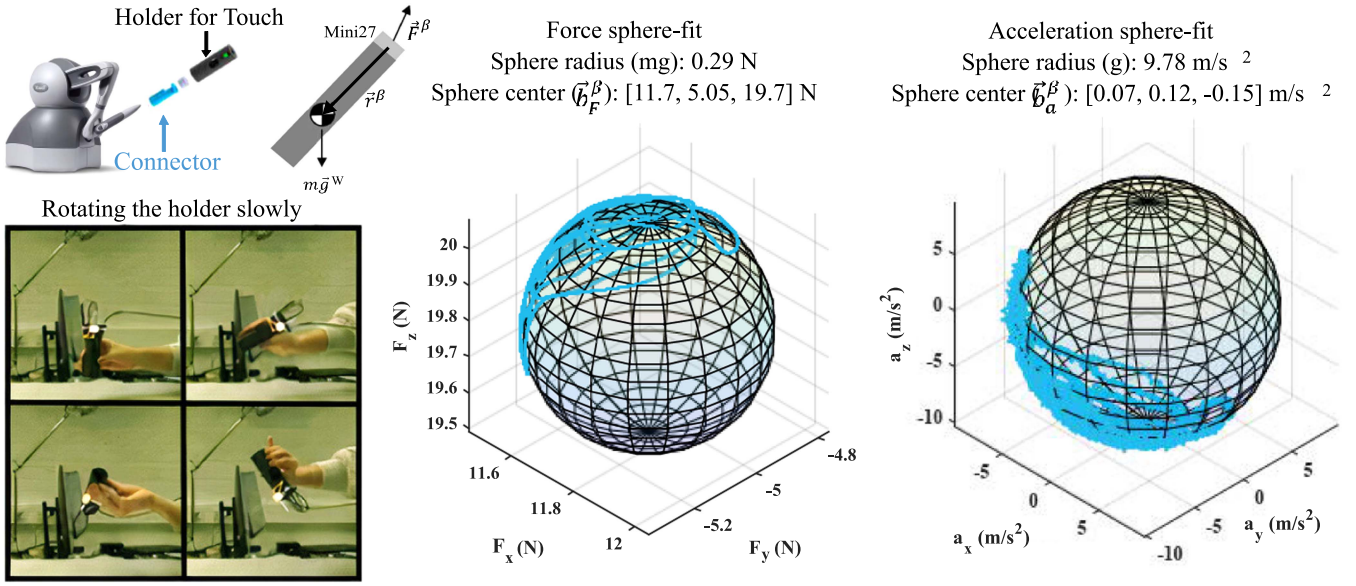

Fig. 3. Calibration process demonstrated for the Touch. To remove the influence of gravity from our readings, we slowly rotate the unattached sensing end-effector in the air. We record force and torque data at 1 kHz with the Mini27, acceleration at 2 kHz with the MPU-9250, and orientation of the end-effector at 100 Hz with the Vicon. The force, torque, and acceleration data are then low-pass filtered, resampled to 100 Hz, and compared with the Vicon data. By calculating the sensor bias (sphere center) and the weight (sphere radius) and center of mass of the connector (3) from these measurements, and by knowing the sensor's world-frame orientation, we can remove the undesired effects of gravity from subsequent interactions recorded by Haptify, leaving only the forces, torques, and vibrations generated by the haptic device being examined.

A. Motion-Capture System Characterization

The Vicon motion-capture system needs a common calibration and origin frame definition before every use. At least three markers are attached to each device link with double-sided tape. Vicon Nexus software is used to record data at 100 Hz. The markers are labeled to define local coordinate frames on each link, and the software provides the transformation matrix from each link's local frame to the origin frame. Merriaux et al. [22] previously evaluated the positioning performance of a Vicon motion-capture system similar to ours, finding a mean absolute positioning error of 0.15 mm for static experiments and less than 2 mm for high-speed experiments.

B. Force Plate Characterization

The force applied to the force plate is equal to the sum of the force readings of the four sensors at the corners. The total torque applied around the center of the force plate is the sum of the four torque readings plus the moments of the four force readings. The force plate is characterized through two experiments to calculate its natural frequency and center of pressure (COP), as described in the Appendix. It provides better force-sensing resolution and more accurate COP estimates than commercial force plates with similar size.

C. Sensing End-Effector Characterization

The sensing end-effector consists of two parts, the connector and the handle, as shown for the Touch in Fig. 3. The connector rigidly attaches the force/torque sensor to the handle of the haptic device. It is designed specifically for each device and 3-D printed. The handle is where the examiner holds; it is rigidly attached to the other side of the force/torque sensor and the accelerometer, and it is the same for both devices tested here, the Touch and the Touch X. For devices with different end-effectors, like the spherical handle of the Force Dimension Omega.3, an appropriately shaped handle and a matching connector are designed and 3-D printed.

We want to measure the forces, torques, and vibrations that the device applies to the user; however, the connector itself has a mass whose weight exerts an additional force and torque. A stationary system such as the force plate requires only one calibration measurement to zero its readings. However, the Mini27 rotates with the end-effector; therefore, to calibrate out the influence of gravity from our readings, we must determine the mass and the location of the center of mass of the connector, and we must know the world-frame orientation of the sensor. Similarly, the accelerometer rotates, so the direction in which it senses gravity constantly changes. Adapting a method explained in [23], we recorded position, orientation, force, acceleration, and torque while slowly rotating the unattached sensing end-effector in all directions in free space, as shown in Fig. 3. Under these conditions, the expected force vector measurement in the handle body frame (β) can be described as follows:

$$\vec{F}^{\beta}(t) = \vec{b}_F^{\beta} + m\vec{g}_W^{\beta}(t) \tag{1}$$

where $\vec{b}_F^{\beta}$ is the sensor's internal force bias (constant offset), m is the mass of the connector, and $\vec{g}_W^{\beta}(t)$ is the constant-magnitude gravitational acceleration vector transformed from the world frame to the handle body frame, whose direction varies over time. Similarly, under these conditions, the total acceleration

vector measurement in the body frame is

$$\vec{a}^{\beta}(t) = \vec{b}_a^{\beta} + \vec{g}_W^{\beta}(t) \tag{2}$$

where $\vec{b}_a^{\beta}$ is the sensor's internal acceleration bias (constant offset). The total torque vector measurement in the handle body frame is

$$\vec{\tau}^{\beta}(t) = \vec{b}_{\tau}^{\beta} + \vec{r}^{\beta} \times m\vec{g}_W^{\beta}(t) \tag{3}$$

where $\vec{b}_{\tau}^{\beta}$ is the internal torque bias and $\vec{r}^{\beta} \times m\vec{g}_W^{\beta}(t)$ is the torque contributed by gravity acting on the mass of the connector. Here, $\vec{r}^{\beta}$ is the constant vector from the sensor to the end-effector's center of mass in the handle body frame. (1) and (2) both have the form of a sphere, as shown in Fig. 3. Fitting a sphere to the force data gives the solution for $\vec{b}_F$ as the center and weight mg as the radius. Fitting a sphere to the acceleration data gives the solution for $\vec{b}_a$ as the center and g as the radius. Finally, we apply robust least squares estimation to solve (3) for $\vec{b}_{\tau}$ and $\vec{r}$. The RMS errors of the fit for the Touch handle are 0.008 N, 0.16 $\mathrm{m/s^2}$, 0.0002 Nm and for the Touch X handle are 0.010 N, 0.17 $\mathrm{m/s^2}$, and 0.0003 Nm. All of these errors are negligible. These fits can thus be used to remove the influence of sensor bias, connector weight, and gravity from the sensing end-effector's readings to highlight the haptic sensations caused by the device itself.

IV. Benchmarking Methods

A Touch and a Touch X were tested both passively and actively to validate Haptify. These two devices were chosen since they are widely known and used in haptics. They are similar and have comparable kinematics, but the Touch X is significantly more expensive. Here, we show how to use Haptify to characterize a GFF device's workspace shape, global free space forces and vibrations, local dynamic forces and toques, frictionless surface rendering, and stiffness rendering.

A. Workspace Shape

The points that can be reached by the end-effector are defined as the reachable workspace [10]. Workspace is a fundamental property of a GFF device because different user applications require different workspaces. As also reported in Haptipedia [4], workspaces are usually provided as three values in the x, y, and z-directions, as if the workspace shape is a rectangular solid with the aforementioned width, height, and depth. However, the real workspace of a GFF device is almost never rectangular. The workspaces given in the technical specifications are 16 cm × 12 cm × 7 cm = 1344 cm^3 for the Touch and 16 cm × 12 cm × 12 cm = 2304 cm^3 for the Touch X. These volumes seem to be a subset of the actual device workspace, where performance is deemed subjectively good, but other manufacturers follow different conventions, and different users have different requirements.

The real workspace shape and volume depend on the device's degrees of freedom (DOF), its kinematic structure, the link lengths, and the joint limits, some of which stem from internal collisions. The workspace can be modeled easily for serial and some hybrid structures if the physical dimensions of the links and the joint limits are known [24]. The Denavit–Hartenberg parameters [25] for the Touch [7] and the Touch X [8] are reported in Table II. The two devices have similar hybrid mechanisms, and their workspaces can be modeled by an equivalent serial kinematic chain. For the α_i, a_1, and d_i values shown in Table II, the end-effector position is

$$P_3^0 = \begin{bmatrix} a_2c_1c_2 + a_3c_1c_2c_3 - a_3c_1s_2s_3 \\ a_2s_1c_2 + a_3s_1c_2c_3 - a_3s_1s_2s_3 \\ a_2s_2 + a_3c_2s_3 + a_3s_2c_3 \end{bmatrix} = \begin{bmatrix} x_{ee} \\ y_{ee} \\ z_{ee} \end{bmatrix} \tag{4}$$

where c_i is $\cos(\theta_i)$ and s_i is $\sin(\theta_i)$. This approach assumes a perfect model of the kinematics with infinitely rigid links, flawless joint angle zeroing, and no external collisions. It also requires one to know the exact values for link lengths and joint limits. Workspace modeling is more complex for parallel mechanisms, where conventional methods like Denavit–Hartenberg do not apply. So far, it has been uncommon for device inventors to provide these geometrical attributes in papers and datasheets.

Our proposed solution for finding the reachable workspace of a device is to experimentally record all the points that its end-effector can reach. The workspace shape and volume can be obtained most efficiently by covering all the corners, edges, and surfaces of the workspace boundary. Like a cube, the workspace of a device with three fixed-range joints has 8 corners, 12 edges, and 6 surfaces. The workspace corners are where all joints are at maximum or minimum value. The workspace edges are where two of the joints are at a limit, and the other joint is varying freely. The workspace surfaces are where we have one joint at its maximum or minimum limit, and the other two joints are moving freely.

TABLE II
Denavit–Hartenberg Parameters for Each Device

Device	i	θ_i (°)	α_i (°)	a_i (m)	d_i (m)
Touch	1	$(-56, 56)$	90	0	0
	2	$(0, 101)$	0	0.135	0
	3	$(-140, -10 + 0.85\,\theta_2)$	0	0.135	0
Touch X	1	$(-52, 52)$	90	0	0
	2	$(0, 100)$	0	0.133	0
	3	$(-140, -20 + 0.85\,\theta_2)$	0	0.133	0

The joint limits came from the SensAble haptics software provided by the company. The link lengths came from [7], [8].

B. Global Free-Space Forces

GFF haptic devices are built to impose forces on the user's body at their end point. Most GFF devices are impedance type and are designed to have low moving mass and low friction [1]. When a passive device exerts high forces on the user's hand, high control effort is required to try to compensate for these intrinsic dynamics in the active mode. These unwanted forces are mostly due to physical causes such as gravity, joint friction, and the inertia of rotational components. Canceling weight reduces the maximum output force, and stably canceling friction or inertia is technically difficult. Ideally, the haptic device should impose zero forces on the user, but these interaction forces exist and are impossible to calculate from the currently provided specifications for each device; device manufacturers also do not provide

this information. We thus propose to capture these forces by having an examiner move slowly within the workspace volume while recording the end-effector position, orientation, velocity, acceleration, and interaction forces, as well as interaction forces at the force plate. The average magnitude and direction of the end-effector forces are calculated to show the global free-space force field across each device workspace. We also compare the force magnitude measured at the end-effector and the force plate for a similar range of translational speeds and angular speeds in both devices.

C. Global Free-Space Vibrations

Haptic devices should render free space cleanly. Any nonintentional vibrations output by a device degrade the rendering quality and can distract the user from the haptic interactions taking place. To quantify this aspect of device behavior, we analyze the vibrations measured in the experiment designed to capture global free-space forces (see Section IV-B). The 3-D acceleration is recorded at a sampling rate of 2 kHz. To eliminate the effect of the user's hand motion, we high-pass filter the signals from 10 Hz. The vibrations in the three directions are combined into one vibration signal using the DFT321 function [26]; this transformation is commonly used to estimate human perception of 3-D vibrations. The reduced 1-D vibration signal has the same spectral energy as the original three-axis acceleration. These measured parasitic vibrations are then quantified by the root-mean-square signal magnitude along with the spectral centroid frequency [27], [28].

D. Local Dynamic Forces and Torques

Haptic devices typically have their best dynamic performance at the center of their workspace, far from singularities and workspace edges. To capture the dynamics of a device, we propose to have the examiner move the end-effector back and forth in all directions at the desired location from low to high speed. These fast oscillatory movements involve high accelerations and thus highlight the device's inertial characteristics. We start from the workspace center and record the interaction forces as well as the end-effector motion.

Real haptic systems present nonlinear behavior, several vibration modes, and distributed masses [29]; therefore, we fit a nonlinear Hammerstein–Wiener (H–W) local dynamic force and torque model to the recorded data using the MATLAB System Identification Toolbox. A data-driven model was chosen since most of the existing analytical models do not consider friction [9], [30], except [31], [32]. In addition, these analytical joint-space models are complex and have strong assumptions about device structure and behavior, so they are impractical to fit to the wide variety of GFF devices we want to benchmark with Haptify. Deriving analytical models for hybrid and parallel mechanisms with friction would be particularly time consuming and system dependent. On the other hand, a data-driven block-oriented model has a simple modeling architecture, a straightforward training method, and a low cost of identification [33]. H–W models are also appropriate for control design and nonlinear systems because they provide flexible parameterization [34],

Fig. 4. Block diagram of a Hammerstein–Wiener nonlinear model.

[35]. As demonstrated in Fig. 4, the H–W model is made of three series blocks. The first block is a nonlinear function *f* that transforms the input data $u(t)$ to $w(t)$, with a similar size

$$w(t) = f(u(t)) \tag{5}$$

The second block is a linear transfer function that transforms $w(t)$ to $x(t)$, with the same dimension

$$x(t) = \frac{B_{ji}}{F_{ji}} w(t). \tag{6}$$

Here, B and F are similar to polynomials in the linear output-error model, and

$$i = 1, 2, \ldots, n_u \text{ and } j = 1, 2, \ldots, n_y \tag{7}$$

where n_u is the number of inputs and n_y is the number of outputs. The last block is a nonlinear function that maps the linear block's output to the system output

$$y(t) = h(x(t)). \tag{8}$$

The 13 inputs for our model are 3-D handle orientation represented as a quaternion plus the 3-D Cartesian position, velocity, and acceleration, all expressed in the world frame. The handle orientation and gimbal-center position and velocity come from the Vicon motion-capture system, and the acceleration comes from the handle-mounted accelerometer after its bias and gravity have been removed. The six model outputs come from the handle force/torque sensor after removal of bias and connector weight. This black-box modeling approach can be used for all GFF devices regardless of their structure or type. It can handle all types of dynamics and global free-space forces, including gravity, joint friction, and nonlinearities. The resulting model can be used by other researchers and future users of the device to predict the forces and torques that will be imposed on the user along recorded or defined trajectories without dealing with the complexities and deficiencies of analytical models.

E. Frictionless Surface Rendering

In an ideal haptic rendering, the user should feel only the interaction forces that originate from the virtual objects being explored [3]. To measure how well each device performs against this ideal, we create a virtual stiff frictionless surface. For simplicity, it is horizontal and located near the center of the device workspace. Ideally, the planar forces felt when moving along this surface should be zero (as they are commanded); any nonzero planar forces stem from the device and not the content being rendered. We expect these forces to be similar to the global free-space forces in the chosen region, though actively exerting normal forces may change how the device feels in the perpendicular directions.

F. Stiffness Rendering

Rendering stiffness is one of the most important capabilities of GFF devices [3]. To examine each device's stiffness rendering quality, we use the same frictionless virtual surface already described for the previous benchmark. We assign it a high constant stiffness value in the vertical (z) direction. Ideally, the normal force should be zero before touching the surface, and it should start increasing precisely when the end-effector reaches the surface. The normal force should then increase linearly with penetration distance until reaching the peak force possible for the device in that configuration.

V. Data Collection

For each GFF haptic device, we conducted five experiments three times each. In all the experiments, the force data were recorded at 1 kHz, the accelerations were recorded at 2 kHz, and the motion data were recorded at 100 Hz.

A. Workspace Experiment

To find the real shape and volume of the workspace, we placed two markers on the end-effector, spanning the center of the gimbal. The haptic device was then placed on a table while the device end-effector position was tracked with the Vicon motion-capture system. The experimenter moved the end-effector to the joint limits and tried to cover all the corners, edges, and surfaces, as explained in Section IV-A. For comparison, each device's theoretical workspace was calculated from its forward kinematics and joint limits.

B. Global Free-Space Experiment

In this experiment, we first took the force plate's bias and calibrated the end-effector's force sensor and accelerometer. We then connected the sensing end-effector to the chosen haptic device and placed it at the center of the force plate. Double-sided tape was used to prevent any unwanted movements of the device's base. The user held the end-effector and moved slowly with a speed less than 0.16 m/s and an angular speed of less than 16 deg/s throughout the workspace, while the device was completely turned off, being careful not to touch the workspace boundaries, so that only free-space forces were captured. The user employed a natural hand pose and orientation through the experiment.

C. Local Dynamic Experiment

In this experiment, after setting up each device as explained in Section V-B, the user performed oscillatory motions in all the directions (x, y, z, and diagonals) with low to high speed for about 140 s. She started near the center of the workspace and was careful not to touch the boundaries of the workspace. No limits were imposed on speed or angular speed; instead, the user started with slow oscillations (around 0.5 Hz) and progressed to motions that were as fast as possible (around 4 Hz), always using a natural hand orientation.

D. Frictionless Surface Rendering

After setting up each device as explained in Section V-B, we connected it via USB 2.0 to a separate computer running the open-source haptic library CHAI3D [36] in Windows 10. A stiff horizontal surface adapted from a CHAI3D example was programmed to have zero friction and a stiffness of 0.75 N/mm. For the experiment, the user pressed the device end-effector lightly into this virtual surface and performed circular motions from low to high speed in the xy plane, centered in the workspace. The commanded data from CHAI3D were recorded at 1 kHz and synchronized with other data streams using the multicontact switch.

E. Stiffness Rendering

This experiment uses the same set up and virtual environment as the frictionless surface rendering benchmark (see Section V-D). The user holds the end-effector near the center of the workspace and repeatedly presses down (in the z direction) into the 0.75 N/mm surface until they reach the edge of the device workspace or the device abruptly stops outputting forces.

VI. Results

Here, we present the results of applying our six benchmarking methods to the data collected in Section V.

A. Workspace Shape

The modeled and measured workspaces for the two devices are presented in Fig. 5. The volume of the measured workspace for the Touch is 11 799 cm^3, while the volume of its model is 13 223 cm^3. The Touch's real reachable workspace of Touch is thus 10.76% smaller than the modeled workspace. This difference mostly comes from the fact that the shape of its base intersects with the modeled workspace and therefore prevents the end-effector from reaching. In addition, the link lengths used in the model cannot be accurately measured without disassembling the device. The volume of the measured workspace for the Touch X is 10 418 cm^3, while the volume of its model is 10 279 cm^3. The reachable workspace of the Touch X is 1.35% bigger than its modeled workspace. Unlike the Touch, the Touch X's base does not interfere with the reachable workspace. More generally, the workspace volumes we measured for the Touch and the Touch X are, respectively, 877% and 452% larger than the rectangular workspaces reported by the manufacturer.

B. Global Free-Space Forces

Fig. 6 shows the measured speeds along with the forces felt while the experimenter moved the end-effectors of the two devices. Both tests had an average translational speed of about 0.05 m/s and an average angular speed of about 4.0 deg/s, which are both slow for the human hand. The total device weights of the Touch and the Touch X are 18.5 N and 30.8 N, respectively. The mean values of -17.4 N, and -30.2 N were subtracted from the force-plate forces in the z direction to facilitate comparisons between the two readings; these values are equal to the weights

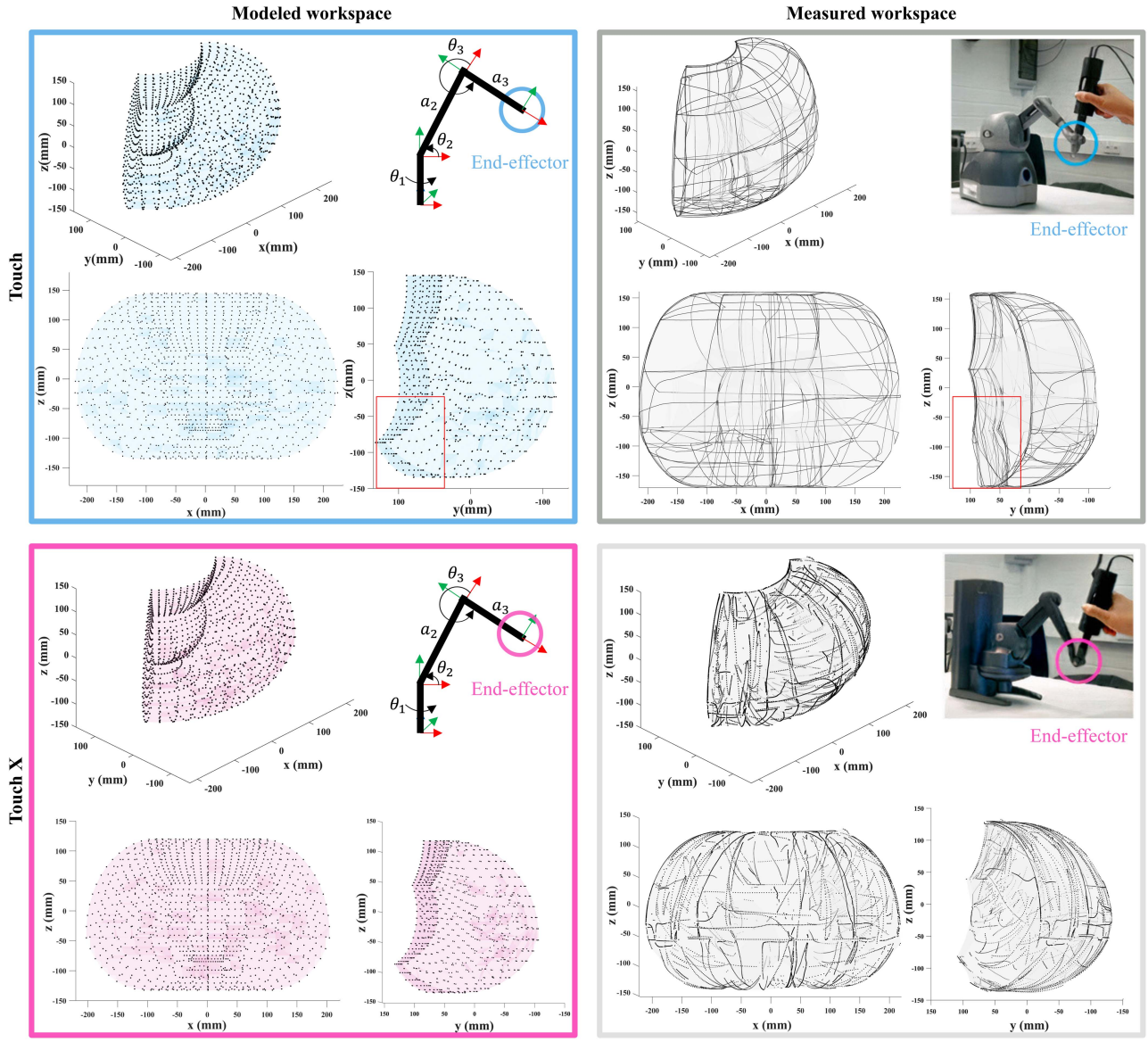

Fig. 5. Modeled and measured workspaces for the Touch and the Touch X. As indicated by the red boxes, the measured workspace of the Touch is missing the lower lobe due to collisions between the end-effector tip and the base.

of the parts of the haptic device that are on the force plate and not held up by the user. The force magnitudes imposed on the user while interacting with the devices vary from 0.13 N to 2.39 N for the Touch and from 0.38 N to 1.18 N for the Touch X. The mean $\pm$ standard deviation of the forces felt by the user are 0.98 N $\pm$ 0.42 N and 0.87 N $\pm$ 0.12 N, respectively.

C. Global Free-Space Vibrations

The high-pass filtered and combined vibrations felt by the user in the global free-space experiments are shown in Fig. 7. The RMS value and the spectral centroid of the combined signal for the Touch are 0.0685 $\mathrm{m/s^2}$ and 122.96 Hz, respectively. The same values for the Touch X are 0.0408 $\mathrm{m/s^2}$ and 162.81 Hz. For comparison, the vibrations measured by Haptify in a similar experiment without any device have an RMS of 0.0332 $\mathrm{m/s^2}$ and a spectral centroid of 108.25 Hz, indicating that both devices cause measurable vibrations. The vibrations recorded from the two devices are also presented as sound files in the supplementary materials. One can listen to or feel these signals to compare the vibration levels between the devices. To create the sound files, we first resampled the vibration signals to a frequency of 44 100 Hz, which is the standard sampling rate for many audio file formats [37]. Second, we divided both signals by the maximum absolute value of the Touch vibration signal (0.8443 $\mathrm{m/s^2}$) to scale both sound files to the maximum of the Touch signal.

D. Local Dynamic Force and Torque Model

The local dynamic experiment data and the outputs of the fitted H–W model for each device are shown in Fig. 8. Each nonlinear H–W model predicts the measured forces and torques in each direction from the measured 3-D position, velocity,

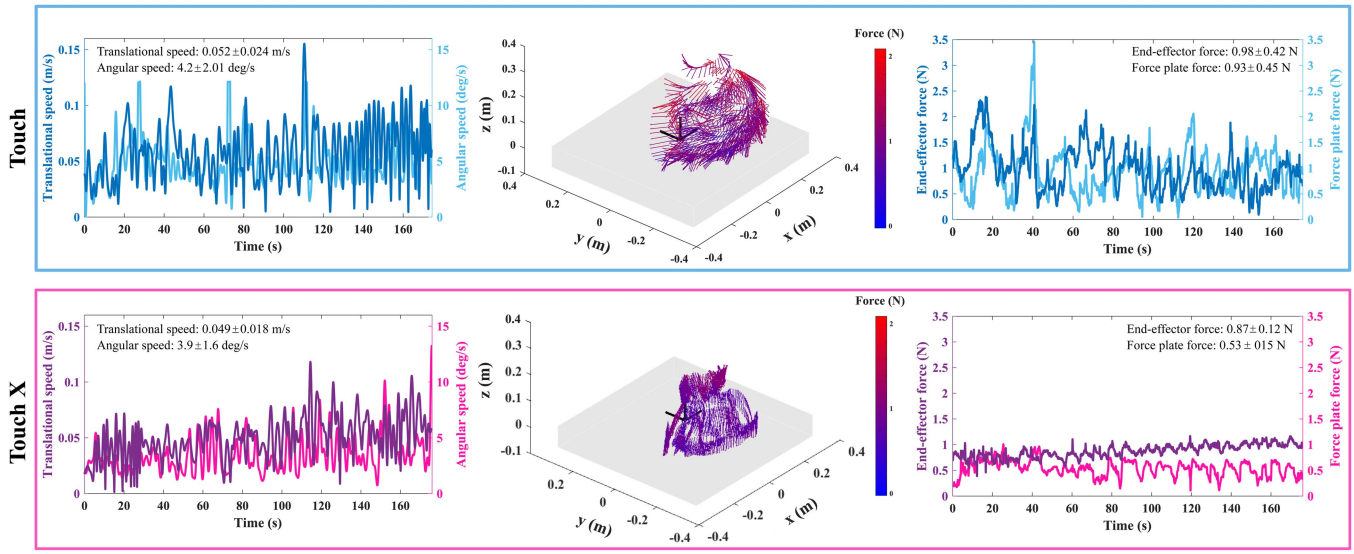

Fig. 6. Global free-space experiments. The experimenter slowly moved the end-effector to a wide range of poses in the workspace while avoiding the workspace edges. The center panel visualizes the force vector measured at each tested position. As shown in the left panel, translational speed averaged about 0.05 m/s, and angular speed averaged about 4.0 deg/s. As seen in the right panel, although the motions are similar, the forces exerted by the Touch on the user at the end-effector are higher and more varied than those exerted by the Touch X. The forces measured by the force plate differ in the same way.

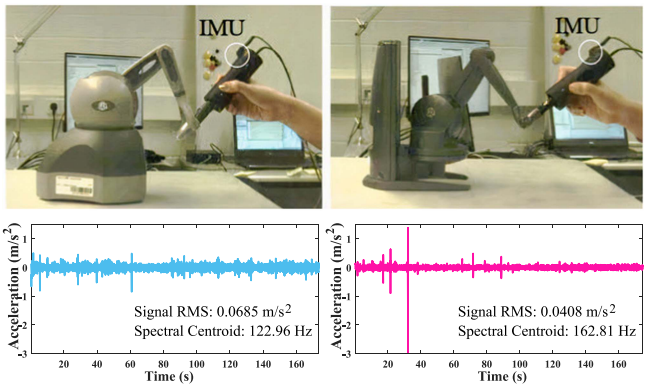

Fig. 7. High-pass filtered and combined vibration signals felt by the user in the global free-space experiments. The signal RMS and spectral centroid are presented in each subfigure; the parasitic vibrations of the Touch (left) are larger and lower frequency than those of the Touch X (right).

acceleration, and orientation (quaternions). Before modeling, the torque readings from the sensor were transferred to the gimbal by subtracting the moment of the measured force vector around the sensor.

With a similar fitting approach and the same number of parameters, the trained model fits the Touch local dynamic experiment with an RMS error of 0.33 N, 0.37 N, and 0.37 N, 0.02 N·m, 0.02 N ·m, and 0.01 N·m for F_x, F_y, F_z, τ_x, τ_y, and τ_z, respectively. The corresponding RMS error values for the Touch X's model are smaller at 0.20 N, 0.13 N, and 0.22 N, 0.005 N · m, 0.004 N·m, and 0.005 N·m. We could reduce the fitting errors by increasing the model complexity; however, such models risk overfitting and tend to incorrectly predict very large forces and torques when tested with unseen trajectories.

To examine and validate the fitted models, we defined a distance metric between the trajectory data on which the models were fit (the training set) and four other experiments that we recorded from the passive devices (test sets). We hypothesized that a good distance metric would be able to predict the error between the modeled and measured forces and therefore show whether a particular model should be trusted to predict what a user would feel when moving a device in a particular way. We began this process by low-pass filtering the velocity data calculated from the Vicon measurements and the accelerometer to eliminate noise and device vibrations. We then subdivided each test set into nonoverlapping time windows that are 5 s long so that the human motion each segment contains is reasonably uniform. Test sets 1 and 2 are from the second and third repetitions of the local dynamic experiment, so they are reasonably similar to the training set. Test set 3 is from an experiment, where the user moved the end-effector in an infinity-shape trajectory with various speeds near the center of the device workspace. Test set 4 is from the global free-space experiment described in Section V-B; it is slow movements over a large area are rather different from the training set.

Each 5-s-long test segment and the whole training set are then further divided into window lengths of 0.2 s to represent the captured motion over a short time period. For each window of data in the test set, we find the most kinematically similar window in the training set by finding the window that minimizes the following distance function:

$$\Delta = \frac{\delta_p}{R_p} + \frac{\delta_v}{R_v} + \frac{\delta_a}{R_a} + \frac{\delta_\theta}{R_\theta}. \tag{9}$$

Here, δ_p, δ_v, δ_a, and δ_θ are the position, velocity, acceleration, and orientation distances between the two selected windows. We define δ_p as the Euclidean distance between the mean positions in the two windows. δ_v is defined as the magnitude of the difference between the average velocity vectors in the two segments, and δ_a is defined the same way for acceleration, as measured by the

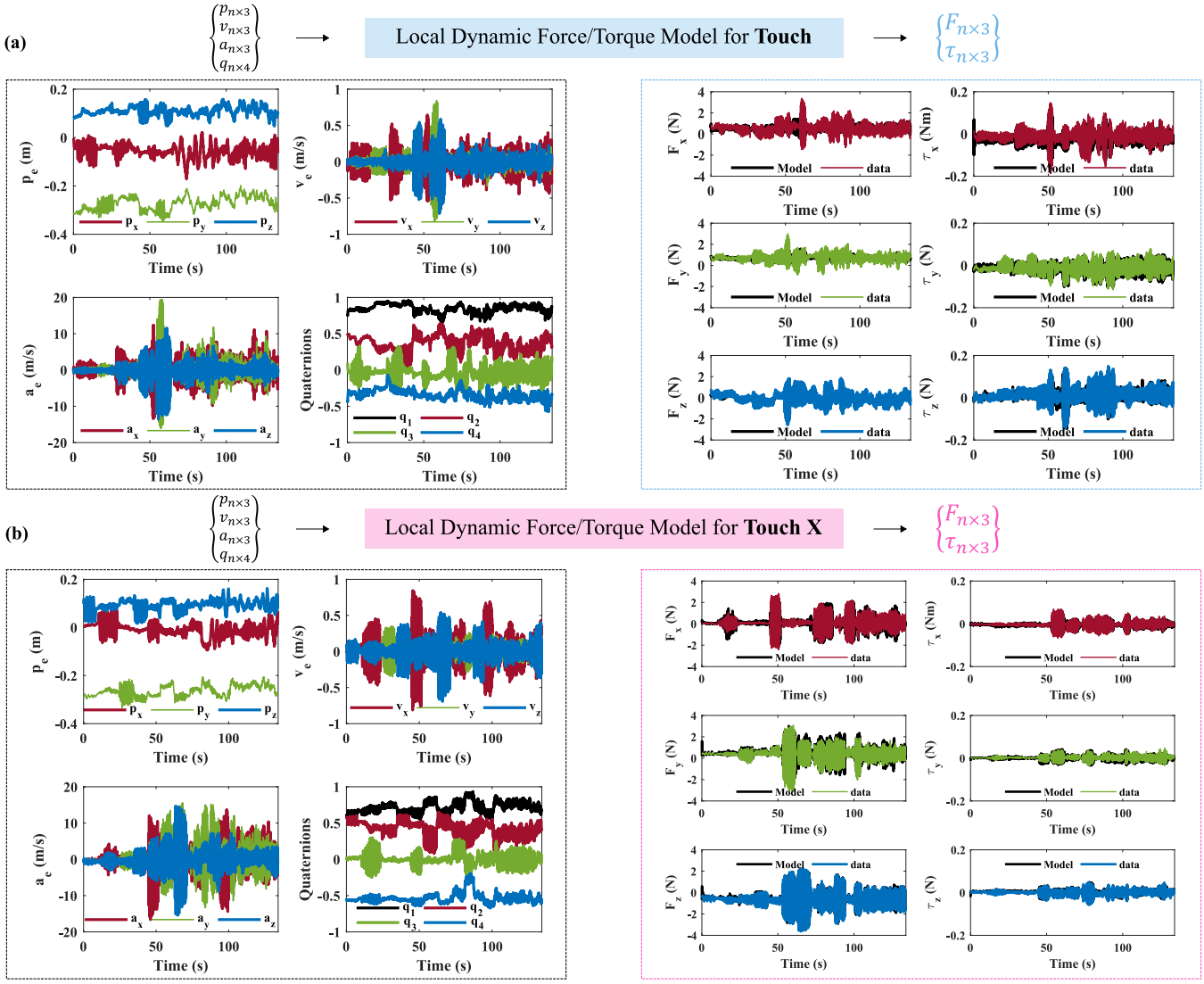

Fig. 8. Trained local dynamic force H–W model for (a) the Touch and (b) the Touch X. The end-effector's position, velocity, acceleration, and orientation are given as inputs (left), and the end-effector forces and torques are the outputs (right). The models capture the recorded forces and torques well.

handle accelerometer. δ_θ is the angle needed to rotate from the average frame orientation in one segment to the average frame orientation in the other segment. These distance values are thus all positive; since they have diverse units, they should not simply be added together. Instead, we normalize them by dividing each measure by a representative distance that we empirically defined to be the maximum difference seen between windows when comparing test sets 1 and 2 with the training set: $R_p = 0.12$ m, $R_v = 1.6\,\mathrm{m/s^2}$, $R_a = 38\,\mathrm{m/s^2}$, and $R_\theta = 86$ deg. We define the overall distance between a particular 5 s long test segment and the training set to be the average of the minimum Δ distances for its 25 windows. In this way, we can calculate a unitless distance that aims to show how different the demonstrated motion is from the motions contained in the training set of a particular local dynamic model.

To quantify the error between the forces predicted by the model and the end-effector forces actually measured during data collection, we passed the 3-D position, velocity, acceleration, and orientation of each test trajectory into the model and calculated the output forces; a similar analysis could be done for output torques. As shown Fig. 9, segments with lower distances tend to have lower force errors. Many segments with a distance less than one have an RMS force error that is comparable to that of the segments from the training set, which have a distance of zero. As anticipated, test sets 1 and 2 have lower distance metrics and lower force errors than test set 4, which is the global free-space experiment. Though the overall motion was different, the segments from test set 3 have relatively low distance metrics and low force errors, indicating successful generalization. As the kinematic distance between a segment and the training set increases, higher force errors are more likely to occur. The black lines show linear models fitted to all presented data points. The model for the Touch has higher force errors than the model for the Touch X, particularly for the slow movements of test set 4. All fitted H–W models are provided as MATLAB files in the supplementary materials. Researchers can use these models

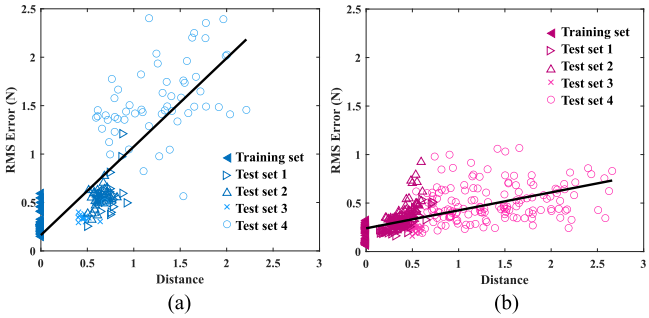

Fig. 9. RMS force prediction error plotted against the kinematic distance metric defined in (9) for (a) the Touch and (b) the Touch X. Each point represents 5 s of data. The training set, test set 1, and test set 2 are three repetitions of the local dynamic experiment. Test set 3 is an infinity-shaped trajectory with varying speed, and test set 4 is the global free-space experiment. As the distance increases, the models tend to have higher force errors, particularly for the Touch.

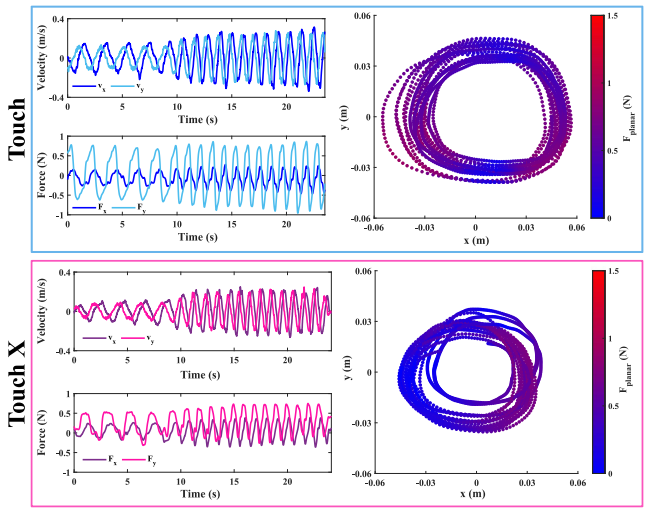

Fig. 10. Planar forces imposed by the devices in the frictionless surface rendering experiment. The Touch imposes higher forces that the Touch X even though the commanded forces were zero for both devices.

to estimate the forces imposed by these two haptic devices for arbitrary recorded or generated trajectories that are inspired by their target application. When a tested trajectory is similar to our training set, the estimated forces should be reasonable. When the distance is high, the model should not be expected to output good predictions.

E. Frictionless Surface Rendering

The planar forces felt during the frictionless surface rendering experiment are shown in Fig. 10. The forces commanded by CHAI3D are zero throughout this experiment, and the experimenter made similar motions. However, the average planar forces imposed by the Touch are 0.53 N, which is higher than the 0.34 N imposed by the Touch X. The Touch's forces are especially large when the end-effector moves in the positive or negative y direction at points farther away from the device midline. The planar forces for the Touch X have lower values overall with only one area of increased magnitude.

F. Stiffness Rendering

Fig. 11 shows the results of the stiffness rendering experiment. The measured stiffness for the Touch is 0.36 N/mm, which is about 47% of the commanded stiffness (0.75 N/mm). The measured stiffness for the Touch X is 0.60 N/mm, or about 80% of the commanded stiffness. As the user continues penetrating into the stiff plane, the normal force imposed by the Touch increases until about 2.7 N. The force remains constant after that point until the user reaches the edge of the workspace. The Touch X shows a different behavior. The force keeps increasing with penetration until about 5.7 N. After that, the force does not remain constant but instead suddenly decreases, causing the user to move very quickly through the stiff surface to the workspace edge. For both devices, the force commanded by CHAI3D keeps linearly increasing with penetration distance.

VII. Discussion and Future Work

A. Workspace Shape

Our results show that the Touch has a larger reachable workspace than the Touch X. The two devices have a similar hybrid structure with slightly different link lengths and joint limits. The Touch's base acts as a rigid constraint and does not let the user reach some locations that are considered to be part of the reachable workspace in the model. Unlike the Touch, the Touch X does not have any collisions inside its workspace, and all the points are reachable. The fact that the manufacturer reports much smaller but differently sized rectangular workspaces for these two devices implicitly shows that device behavior varies across the workspace, foreshadowing the need to investigate both passive and active device performance as a function of position. In future extensions of this work, we also plan to use measurements to obtain the device's kinematic chain, joint limits, link lengths, rotational workspace, and manipulability.

B. Global Free-Space Forces

The forces imposed on the user's hand in the global free-space experiment with the passive Touch are higher than those from the passive Touch X; this result matches what the user perceived during the experiments. As shown in Haptipedia [4], the manufacturer reported higher translational friction for the Touch (0.26 N) than for the Touch X (0.06 N), though both of these values are much less than what we measured; our higher force measurements most likely come from the lack of gravity compensation. In future work, we plan to try to distinguish the effects of gravitational forces from friction, rotational inertia, and other sources in the global free-space forces. We also intend to evaluate the software-based gravity compensation provided by some device manufacturers and investigate automatic data-driven compensation methods.

C. Global Free-Space Vibrations

The combined vibration signal in the interaction with the Touch is stronger than the Touch X. The spectral centroid is higher for the Touch X, meaning its vibrations have higher

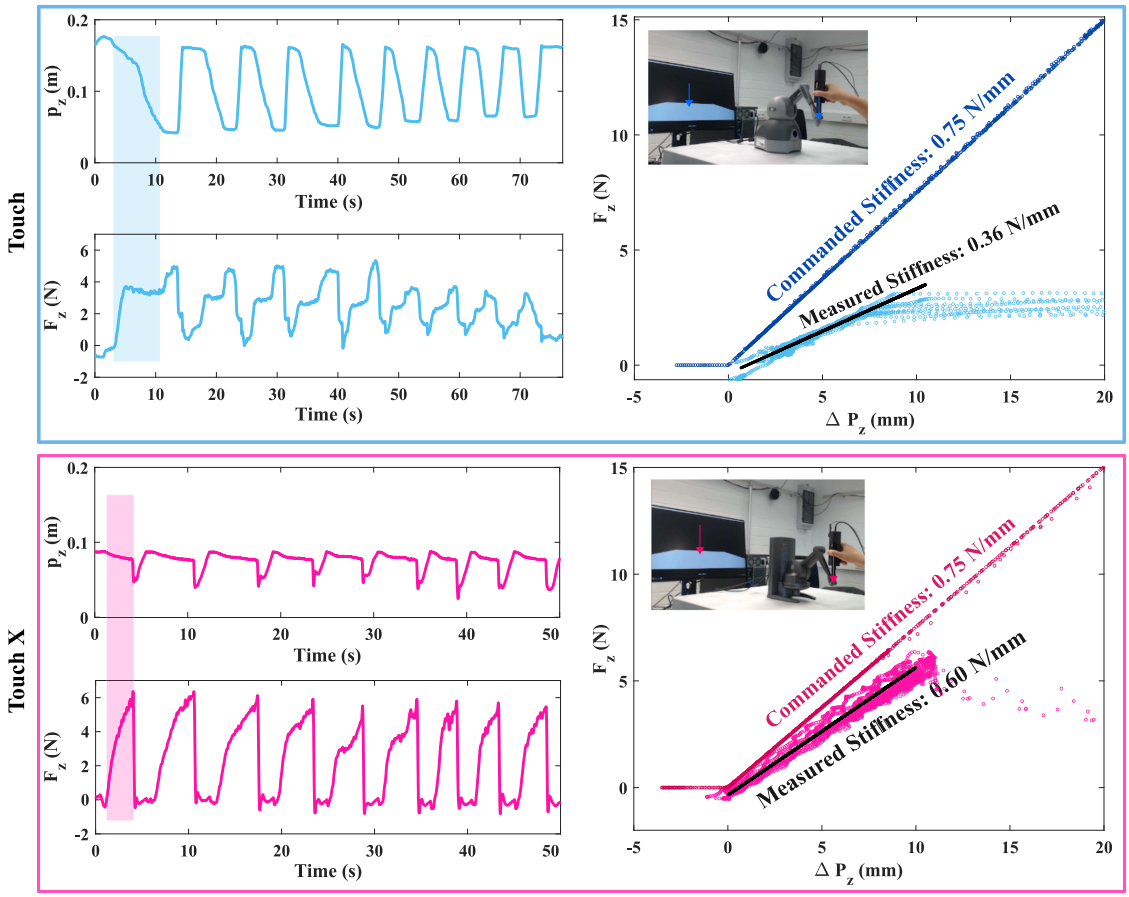

Fig. 11. Stiffness rendering experiment. F_z is the force applied to the user's hand in the z direction, p_z is the z position of the end-effector, and ΔP_z is the penetration of the end-effector into the virtual surface. The commanded stiffness is 0.75 N/mm, but neither device accurately renders this stiffness. The measured stiffness for the Touch is 0.36 N/mm and for the Touch X is 0.60 N/mm.

pitch, on average. The different levels of vibrations in these two devices with similar kinematic structures could be due to the different tolerance of bearings, different stiffness of segments, backlash in the drive train, the passive behavior of the motors, and loose or tight tolerances on joint axles. Although users may complain about parasitic vibrations in a particular GFF device, such measurements have not previously been proposed as an indicator of quality.

D. Local Dynamic Force and Torque Model

The H–W models fitted for both devices show good predictions for input trajectories that are similar to the trajectories on which the models were trained, as quantified by our distance metric. With relatively similar input motion and the same modeling approach, the Touch X model estimates forces for unseen trajectories with lower errors than the Touch model. The fitting errors achieved during training are also lower for the Touch X, meaning its model more accurately captures the measured behavior. We believe that the higher mass of the links, higher joint friction, cable flexibility, backlash, and other nonlinear effects make the Touch's model less accurate. We also observe that the Touch forces are larger and vary more with position in the global free-space experiment. Therefore, it is not surprising that force prediction is less accurate for trajectories farther from the center area, where the models were trained. One way of improving the prediction performance would be to train a model for a large range of trajectories at various locations in the workspace, though such recordings require more time to gather and to process. In future work, we plan to investigate the utility of using deduced joint angle values as direct inputs to the model. We will also test how well these approaches apply to other devices with more distinct mechanisms. Modeling the feel of devices with low DOF will be particularly interesting, as workspace limits cannot be avoided.

E. Frictionless Surface Rendering

Although the commanded planar forces were zero, both devices imposed nonzero planar forces on the user's hand, and the Touch exerted higher forces than the Touch X. We believe these forces mainly stem from the friction and inertia of the device itself, which matches with the results of the passive global free-space experiment. It appears that these two devices do not compensate for gravity and friction when actively rendering virtual content with CHAI3D; therefore, the lower quality of

TABLE III
HAPTIFY BENCHMARKS FOR THE TWO EVALUATED GFF DEVICES

Metric	Touch	Touch X
Workspace Volume (cm^3)	11799	10418
Mean Free-Space Forces (N)	0.98	0.87
Free-space Vibration RMS Magnitude (m/s^2)	0.0685	0.0408
Free-space Vibration Spectral Centroid (Hz)	122.96	162.81
Dynamic Model Mean Force Estimation Error (N)	0.35	0.18
Dynamic Model Mean Torque Estimation Error (Nm)	0.016	0.004
Frictionless Surface Rendering Forces (N)	0.53	0.38
Stiffness Rendering Accuracy (%)	47.3	80
Stiffness Peak Force (N)	2.7	5.7

free space in the Touch leads to lower quality of rendering a frictionless surface. In future work, we plan to investigate the connection between these two benchmarks and link our measures to human perception.

F. Stiffness Rendering

In line with our other results, the Touch X renders high stiffness more accurately than the Touch, though neither device produces an end-effector interaction that matches the commanded stiffness. We believe the main cause for this softening is the series elasticity of the device drive train; indeed, the Touch's internal cable drive system includes springs that stretch when active forces are exerted. The peak normal force measured for the Touch and the Touch X in this experiment are 2.7 and 5.7 N, while the manufacturer reports peak forces of 3.3 and 7.9 N, respectively. The discrepancy likely comes from the uncompensated effects of gravity and/or the direction in which the force is applied. The distance between the end-effector and the device base also affects the stiffness rendering quality and the maximum force available. We have anecdotally observed more vibrations when pressing down far from the base, especially when higher forces are commanded. The speed and acceleration of the end-effector also affect the measured forces. The experimenter tried to keep the motion similar across trials, but there were naturally some differences, which likely caused some of the data scatter seen in Fig. 11. Future work needs to investigate rendering of a range of stiffness values, the uniformity of stiffness across the workspace, and rendering of other haptic effects such as viscous damping and textures.

G. Design of Haptify

We designed and built Haptify to address the need for a standard benchmarking system for GFF haptic devices. Our results have proven that attaching external motion, force, and acceleration sensors to a device enables measurement of haptic attributes that are not provided by the manufacturer and that cannot easily be modeled or calculated. Our proposed benchmarks are shown in Table III. Now that we have validated Haptify with experiments and analysis on two commercial haptic interfaces, we hope to benchmark a diverse set of GFF devices.

To have the highest quality as a standard benchmark for a large selection of devices, Haptify uses expensive, robust components that are not portable. It requires a calibration process to assure accurate recordings. Future extensions of our work will include a more detailed analysis of the recordings along with more sophisticated investigations of active haptic rendering. We also plan to run a user study to find out the techniques that haptic experts use to compare device characteristics in both passive and active modes. We plan to make our findings from Haptify accessible to all researchers by adding extra features to the Haptipedia website. These detailed, reliable, and relevant measures should help designers efficiently choose devices that match their task requirements. By showing specific device attributes, Haptify could inspire device designers and engineers to create new devices with better performance. Haptify's evaluation of GFF devices may also lead to hardware changes and software approaches to compensate for the documented nonidealities, which may enable existing devices to achieve higher performance.

VIII. CONCLUSION

This article presented Haptify, a comprehensive benchmarking system for GFF haptic devices. Haptify has three main sensing components: a motion-capture system, a custom-built force plate, and a sensing end-effector. The sensors used in Haptify were carefully chosen to be usable for many devices and cover the limits of most GFF devices' output forces, torques, and position ranges. We validated Haptify by evaluating two common commercial haptic devices and showed that Haptify's external sensors allow us to characterize the two devices in a standard, noninvasive way.

We compared each device's measured workspace with its modeled workspace and showed how imprecise joint limits, incorrect link lengths, and unmodeled collisions affect the workspace's shape and volume. To characterize the global free-space forces and vibrations, we recorded the interaction forces, end-effector position, orientation, velocity, and acceleration while the user moved slowly through the workspace. Our results showed that the Touch exerts higher free-space forces and larger vibrations on the user than the Touch X does, confirming the superiority of the more expensive device. For the local dynamic force and torque model, the user moved from low to high speeds in the x, y, z, and diagonal directions. We validated our local dynamic models by comparing the estimated forces with the recorded forces of other experiments and showed that the force error is small for trajectories similar to the training set. This H–W modeling approach can be used for all future haptic devices evaluated by Haptify, regardless of their structure type, DOF, and complexities. Furthermore, the distance metric we developed should enable researchers to determine whether a given model is likely to accurately predict the forces for a new trajectory of interest. We evaluated the active rendering performance of the devices via two experiments with a frictionless stiff surface. As expected, the passive device properties strongly affected active device performance, with the Touch showing higher planar forces than the Touch X. Most surprisingly, we found that the stiffness felt by the user is significantly lower than the commanded stiffness for both devices.

We have four main directions for future work. The first direction is to define and measure more benchmarking methods, particularly for active haptic rendering. We plan to conduct a

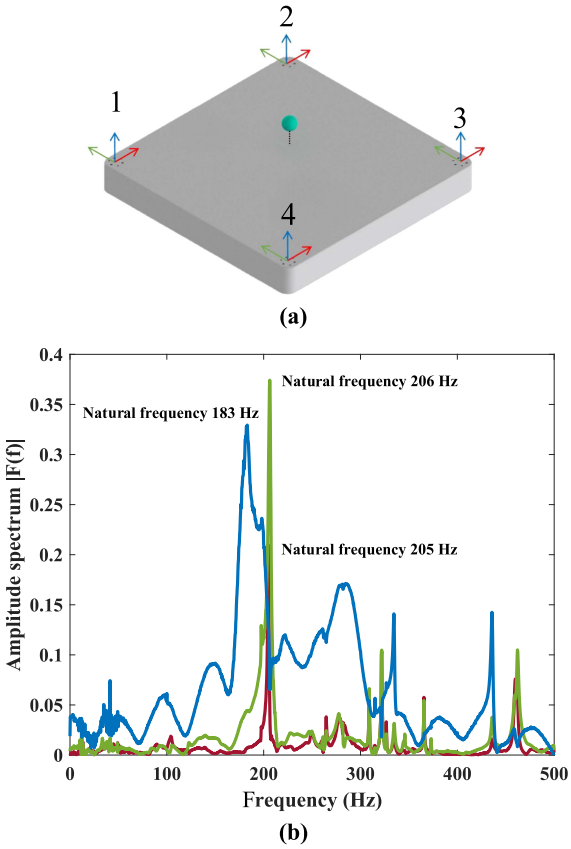

Fig. 12. (a) The impact test in which a ball is dropped at the center of the force plate. (b) The corresponding frequency responses. F_x shows a prominent peak at 205 Hz, F_y shows a prominent peak at 206 Hz, and F_z shows a prominent peak at 182 Hz.

user study wherein haptic experts try different haptic devices as we record the interactions with Haptify. These experiments should allow us to further elucidate parameters that define a good active haptic device and a good haptic interaction. Once we have standardized a complete set of benchmarks, we intend to create a graphical user interface to guide the experimenter through each test, which should increase the quality of the resulting metrics and reduce the experimenter's mental workload [38]. Our third direction is to benchmark more devices and make our findings and measurements accessible for all researchers by adding features to the Haptipedia website. The fourth direction is to build a less expensive and yet sufficiently precise version of Haptify.

APPENDIX

1) Natural Frequency Experiment: The natural frequency of the force plate was determined through an impact test by dropping a ball, as described in [39] and shown in Fig. 12(a). The transient response to the impact was measured at 1 kHz. A fast Fourier transform (FFT) was applied to visualize the frequency response of the force plate in each of the three Cartesian directions [see Fig. 12(b)]. Each natural frequency was determined as the most prominent peak in the frequency response. All are around 200 Hz, which is far beyond the bandwidth of human motion.

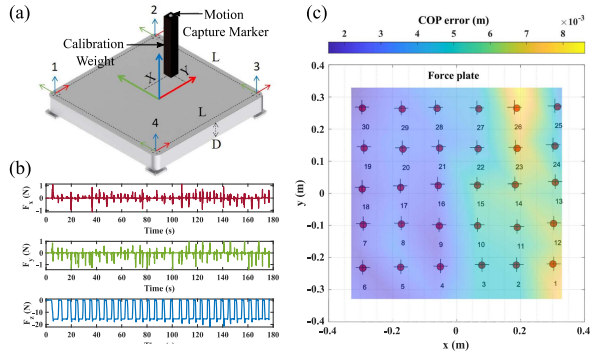

Fig. 13. (a) The experiment process, where a weight was placed in thirty different locations on the force plate. (b) The force data recorded at 1 kHz by the force plate. (c) The results; red circles represent the COP calculated from the force plate measurements, and black plus signs are the COP measured from the Vicon data.

2) COP Experiment: We calculated the COP for the force plate based on Kistler documentation [40]. A weight was placed in 30 equally spaced locations. X and Y, the coordinates of the object with respect to the center force plate, are calculated as $X = \frac{-M_y}{F_z}$ and $Y = \frac{-M_x}{F_z}$, where $M_x = \frac{L}{2}(F_{1z} + F_{2z} - F_{3z} - F_{4z})$, $My = \frac{L}{2}(-F_{1z} + F_{2z} + F_{3z} - F_{4z})$, and the center-to-center distance between sensors is $L = 60$ cm. The forces measured in the COP experiment are shown in Fig. 13(b), and the 30 COPs calculated from the force plate readings are plotted as red circles in Fig. 13(c). The error between the COP calculated and the COP measured from the Vicon motion-capture system is 4.8 ± 3.8 mm. Three existing commercial Kistler force plates, 9281B11, 9281 C, and 9287B, with dimensions of 40 cm $\times$ 60 cm, 40 cm $\times$ 60 cm, and 60 cm $\times$ 90 cm are studied in [40] and have COP errors of 8.4 ± 4.4 mm, 11.6 ± 6.1 mm, and 14.1 ± 9.1 mm, respectively [40].

ACKNOWLEDGMENT

The authors would like to thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting Farimah Fazlollahi, the Central Mechanical Workshop for fabricating the aluminum force plate, the MPI-IS Robotics ZWE for 3-D printing and 3-D scanning, Jonathan Fiene for developing the accelerometer board and M2 microcontroller, Hasti Seifi, Karon MacLean, and Alborz Aghamaleki Sarvestani for many fruitful discussions, Bernard Javot and Joey Burns for helping set up the Haptify laboratory, Behnam Khojasteh for input on vibration analysis, and Natalia Sanchez-Tamayo and Neha Thomas for helpful feedback on this manuscript.

REFERENCES

[1] K. J. Kuchenbecker, "Haptics and haptic interfaces," in *Encyclopedia of Robotics*. Berlin, Germany: Springer, 2018.
[2] V. Hayward, O. R. Astley, M. Cruz-Hernandez, D. Grant, and G. Robles-De-La-Torre, "Haptic interfaces and devices," *Sensor Rev.*, vol. 24, no. 1, pp. 16–29, 2004.
[3] T. H. Massie and Salisbury, "The phantom haptic interface: A device for probing virtual objects," in *Proc. ASME Winter Annu. Meeting, Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, Chicago, IL, 1994, pp. 295–300.

[4] H. Seifiet al., "Haptipedia: Accelerating haptic device discovery to support interaction and engineering design," in *Proc. ACM CHI Conf. Hum. Factors Comput. Syst.*, 2019, pp. 1–12.
[5] H. Seifi, M. Oppermann, J. Bullard, K. E. MacLean, and K. J. Kuchenbecker, "Capturing experts' mental models to organize a collection of haptic devices: Affordances outweigh attributes," in *Proc. ACM CHI Conf. Hum. Factors Comput. Syst.*, 2020, pp. 1–12.
[6] A. J. Silva, O. A. D. Ramirez, V. P. Vega, and J. P. O. Oliver, "Phantom omni haptic device: Kinematic and manipulability," in *Proc. IEEE Electron., Robot. Automot. Mechanics Conf.*, 2009, pp. 193–198.
[7] M. H. Koul, P. Kumar, P. K. Singh, M. Manivannan, and S. K. Saha, "Gravity compensation for phantom omni haptic interface," in *Proc. Joint Int. Conf. Multibody Syst. Dyn.*, 2010, pp. 25–27.
[8] Y. Jiang, C. Yang, X. Wang, and C.-Y. Su, "Kinematics modeling of geomagic touch x haptic device based on adaptive parameter identification," in *Proc. IEEE Int. Conf. Real-Time Comput. Robot.*, 2016, pp. 295–300.
[9] M. C. Çavuşoğlu, D. Feygin, and F. Tendick, "A critical study of the mechanical and electrical properties of the phantom haptic interface and improvements for high performance control," *Presence, Teleoperators Virtual Environ.*, vol. 11, no. 6, pp. 555–568, 2002.
[10] E. Samur, *Performance Metrics for Haptic Interfaces*. Berlin, Germany: Springer Science & Business Media, 2012.
[11] C. Hatzfeld and T. A. Kern, *Engineering Haptic Devices*. Berlin, Germany: Springer, 2016.
[12] R. Madhavan, R. Lakaemper, and T. Kalmár-Nagy, "Benchmarking and standardization of intelligent robotic systems," in *Proc. IEEE Int. Conf. Adv. Robot.*, 2009, pp. 1–7.
[13] A. Aly, S. Griffiths, and F. Stramandinoli, "Metrics and benchmarks in human-robot interaction: Recent advances in cognitive robotics," *Cogn. Syst. Res.*, vol. 43, pp. 313–323, 2017.
[14] D. Feil-Seifer, K. Skinner, and M. J. Matarić, "Benchmarks for evaluating socially assistive robotics," *Interact. Stud.*, vol. 8, no. 3, pp. 423–439, 2007.
[15] J. Weisz, Y. Huang, F. Lier, S. Sethumadhavan, and P. Allen, "RoboBench: Towards sustainable robotics system benchmarking," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2016, pp. 3383–3389.
[16] H. Culbertson and K. J. Kuchenbecker, "Importance of matching physical friction, hardness, and texture in creating realistic haptic virtual surfaces," *IEEE Trans. Haptics*, vol. 10, no. 1, pp. 63–74, Jan./Mar. 2017.
[17] K. J. Kuchenbecker and G. Niemeyer, "Modeling induced master motion in force-reflecting teleoperation," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2005, pp. 348–353.
[18] K. J. Kuchenbecker and G. Niemeyer, "Induced master motion in force-reflecting teleoperation," *ASME J. Dyn. Syst., Meas., Control*, vol. 128, no. 4, pp. 800–810, Dec. 2006.
[19] C. Frissonet al., "RepHap: Towards an opensource platform for benchmarking haptic devices leveraging the robot operating system ecosystem," in *Proc. IEEE Haptics Symp.*, 2020, pp. 586–587.
[20] B. Siciliano, L. Sciavicco, L. Villani, and G. Oriolo, *Robotics: Modelling, Planning and Control*. Berlin, Germany: Springer Science & Business Media, 2010.
[21] R. Q. Van der Linde, P. Lammertse, E. Frederiksen, and B. Ruiter, "The HapticMaster, a new high-performance haptic interface," in *Proc. Eurohaptics*, 2002, pp. 1–5.
[22] P. Merriaux, Y. Dupuis, R. Boutteau, P. Vasseur, and X. Savatier, "A study of vicon system positioning performance," *Sensors*, vol. 17, no. 7, 2017, Art. no. 1591.
[23] A. Burkaet al., "Proton: A visuo-haptic data acquisition system for robotic learning of surface properties," in *Proc. IEEE Int. Conf. Multisensor Fusion Integration Intell. Syst.*, 2016, pp. 58–65.
[24] D. Zhang, F. Cursi, and G.-Z. Yang, "WSrender: A workspace analysis and visualization toolbox for robotic manipulator design and verification," *IEEE Robot. Automat. Lett.*, vol. 4, no. 4, pp. 3836–3843, 2019.
[25] M. W. Spong, S. Hutchinson, and M. Vidyasagar, *Robot Modeling and Control*, vol. 3, Hoboken, NJ, USA: Wiley, 2006.
[26] N. Landin, J. M. Romano, W. McMahan, and K. J. Kuchenbecker, "Dimensional reduction of high-frequency accelerations for haptic rendering," in *Proc. Int. Conf. Hum. Haptic Sens. Touch Enabled Comput. Appl.*, 2010, pp. 79–86.
[27] G. J. Sandell, "Roles for spectral centroid and other factors in determining "blended" instrument pairings in orchestration," *Music Percep.*, vol. 13, no. 2, pp. 209–246, 1995.
[28] G. C. Ozmen, M. Safaei, L. Lan, and O. T. Inan, "A novel accelerometer mounting method for sensing performance improvement in acoustic measurements from the knee," *J. Vib. Acoust.*, vol. 143, no. 3, 2020, Art. no. 031006.
[29] I. Díaz and J. J. Gil, "Influence of vibration modes and human operator on the stability of haptic rendering," *IEEE Trans. Robot.*, vol. 26, no. 1, pp. 160–165, Feb. 2009.
[30] H. Kazerooni and M.-G. Her, "The dynamics and control of a haptic interface device," *IEEE Trans. Robot. Automat.*, vol. 10, no. 4, pp. 453–464, Aug. 1994.
[31] A. M. Tahmasebi, B. Taati, F. Mobasser, and K. Hashtrudi-Zaad, "Dynamic parameter identification and analysis of a phantom haptic device," in *Proc. IEEE Conf. Control Appl.*, 2005, pp. 1251–1256.
[32] B. E. Miller, J. E. Colgate, and R. A. Freeman, "On the role of dissipation in haptic systems," *IEEE Trans. Robot.*, vol. 20, no. 4, pp. 768–771, Aug. 2004.
[33] E. Shokrollahi, A. A. Goldenberg, J. M. Drake, K. W. Eastwood, and M. Kang, "Application of a nonlinear Hammerstein-Wiener estimator in the development and control of a magnetorheological fluid haptic device for robotic bone biopsy," *Actuators*, vol. 7, no. 4, 2018, Art. no. 83.
[34] S. I. Biagiola and J. L. Figueroa, "Wiener and Hammerstein uncertain models identification," *Math. Comput. Simul.*, vol. 79, no. 11, pp. 3296–3313, 2009.
[35] J.-W. van Wingerden and M. Verhaegen, "Closed-loop subspace identification of Hammerstein-Wiener models," in *Proc. IEEE Conf. Decis. Control*, 2009, pp. 3637–3642.
[36] 2022. [Online]. Available: https://chai3d.org
[37] J. Watkinson, "Digital audio," *J. Audio Eng. Soc.*, vol. 36, no. 6, pp. 492–508, 1988.
[38] A. Abdulali, I. R. Atadjanov, and S. Jeon, "Visually guided acquisition of contact dynamics and case study in data-driven haptic texture modeling," *IEEE Trans. Haptics*, vol. 13, no. 3, pp. 611–627, Jul./Sep. 2020.
[39] G. Paolini, U. Della Croce, P. O. Riley, F. K. Newton, and D. C. Kerrigan, "Testing of a tri-instrumented-treadmill unit for kinetic analysis of locomotion tasks in static and dynamic loading conditions," *Med. Eng. Phys.*, vol. 29, no. 3, pp. 404–411, 2007.
[40] *Improving COP Accuracy with Kistler Force Plates*, Winterthur, Switzerland: Kistler, 2019.

Farimah Fazlollahi (Graduate Student Member, IEEE) received the B.S. degree in mechanical engineering with a major in applied design and the M.S degree in mechanical engineering with a major in applied design, dynamics, vibrations, and control from Shiraz University, Shiraz, Iran, in 2015 and 2017, respectively. She is currently working toward the Ph.D. degree in mechanical engineering with the Haptic Intelligence Department, Max Planck Institute for Intelligent Systems, Stuttgart, Germany.

Her research interests include haptics, robotics, dynamics, mechanism design, and sensor fusion.

Katherine J. Kuchenbecker (Fellow, IEEE) received the Ph.D. degree in mechanical engineering from Stanford University, Stanford, CA, USA, in 2006.

She did postdoctoral research with the Johns Hopkins University, Baltimore, MD, USA. She directs the Haptic Intelligence Department, Max Planck Institute for Intelligent Systems, Stuttgart, Germany. She was an Engineering Professor with the University of Pennsylvania, Philadelphia, PA, USA, before she moved to the Max Planck Society in 2017. Her research centers on haptic interfaces, which enable a user to touch virtual and distant objects as though they were real and within reach, as well as haptic sensing systems, which allow robots to physically interact with objects and people.

Dr. Kuchenbecker delivered a TEDYouth talk on haptics in 2012 and has been honored with a 2009 NSF CAREER Award, the 2012 IEEE Robotics and Automation Society Academic Early Career Award, a 2014 Penn Lindback Award for Distinguished Teaching, and various paper, demonstration, and reviewing awards. She co-chaired the IEEE Technical Committee on Haptics from 2015 to 2017 and the IEEE Haptics Symposium in 2016 and 2018.