Measurement-based Modeling of Contact Forces and Textures for Haptic Rendering

Jochen Lang, Member, IEEE, and Sheldon Andrews, Member, IEEE

Abstract—Haptic texture represents the fine-grained attributes of an object's surface and is related to physical characteristics such as roughness and stiffness. We introduce an interactive and mobile scanning system for the acquisition and synthesis of haptic textures that consists of a visually tracked hand-held touch probe. The most novel aspect of our work is an estimation method for the contact stiffness of an object based solely on the acceleration and forces measured during stroking of its surface with the hand-held probe. We establish an experimental relationship between the estimated stiffness and the contact stiffness observed during compression. We also measure the height-displacement profile of an object's surface enabling us to generate haptic textures. We show an example of mapping the textures onto a coarse surface mesh obtained with an image-based technique, but the textures may also be combined with coarse surface meshes obtained by manual modeling.

Index Terms—Haptics, Texture, Contact Stiffness, Measurement-based modeling

1 INTRODUCTION

HAPTICS refers to the stimulation of a person's sense of touch due to manual interaction with an object or environment. Through computer controlled forces, haptic technology is able to display tactile and kinesthetic cues arising from interaction in a virtual setting. This exposes characteristics about the application environment that cannot be easily purveyed using graphical or acoustic displays, and unlike these other displays, haptic devices act as both an input and output human-machine interface. This bi-directionalism is a distinguishing feature of haptic interfaces [1], i.e., a simultaneous exchange of information occurs between the machine and the user. Areas of application for this technology are widespread—medicine, entertainment, computer aided design, and other multimedia applications stand to benefit from haptic technology.

Rendering of haptic phenomena is often done based on models that represent physical attributes of objects in the application environment, including stiffness and roughness. It has been shown that detailed surface characteristics have a significant effect on object identification [2], and that in the absence of structural feedback material detail is solely relied on for identification. This indicates that the realism of virtual objects may be increased by assigning realistic properties used in surface interaction.

Assigning values to properties used for haptic rendering can sometimes be done manually, but realistic rendering often requires scanning of real-world object properties. Scanning physical interaction behaviour has previously required a complex and expensive robotic environment [3], [4], which is cumbersome to use. In this paper, we propose a mobile scanning system capable of scanning tactile surface features, in particular roughness and stiffness, which overcomes these limitations. Our system does not require a robotic actuator. Instead, it is implemented using inexpensive and readily available components, which makes it an attractive alternative to other tactile scanning systems. By registering scans with coarse 3D geometry, we show that our system can be employed to haptically texture virtual objects without the need for high-resolution 3D meshes. However, this does not exclude the option of integrating our system as part of the standard 3D scanning pipeline [5].

We have described a preliminary version of our system in the conference 3D-Imaging and Modeling 2007 [6] and our texturing approach in the Workshop on Haptic Audio Virtual Environments and Games 2007 [7]. However, the main focus of this paper is to present a novel approach to obtain a measure of the stiffness of the surface based on scanning and to present, for the first time, an extensive evaluation of our approach based on a complete example of scanning a wooden statue (shown in Figure 1(a)). The performance of our system has been significantly improved compared to our initial implementation due to higher accuracy in visual tracking and a sensor fusion for the scan profile.
Fig. 1. Our scanning system in action. Figure 1(a) shows the user employing the WHaT force sensor to acquire haptic texture and forces. Figures 1(c) and 1(d) show the result of the modeling effort with textured areas highlighted in Figure 1(c).

1.1 Background

Describing textures and reproducing them in a virtual environment is problematic due in part to the fact that it is difficult to say which properties best characterize the surface of an object. Rough-smooth, hard-soft, slippery-sticky, and rigid-elastic are all descriptors for the feel of a surface. A study by Hollins et al. [8] identified the dimensionality of surface texture by asking subjects to categorize 17 texture samples based on the perceived similarity when the samples were stroked with a finger. Their study found that we perceive texture mainly in the dimensions of rough-smooth and hard-soft. Bergmann Tiest and Kappers [9] conducted a similar experiment, with 124 sample materials, for which the physical parameters of compression force related to the material’s bulk-modulus and roughness were measured. Klatzky and Lederman [10], [11] found that when an intermediary object is interposed between a surface and skin (analogous to a stylus-based haptic display), vibrotactile coding determines the perceived roughness of the surface. A study by Pongrac [12] showed that frequency and amplitude are key factors in perception of texture when vibrotactile coding is used. Others studies have shown the sensitivity of mechanoreceptors found in the skin to peak at around 300 Hz [13]. The geometry and physical parameters of a surface are not the only factors affecting perception when considering computer-based haptic texturing. Campion and Hayward [14] showed that perception is also affected by the bandwidth and the electromechanical transfer function of the haptic device and needs to be considered when rendering textures, including the ones acquired by our system. In summary, vibrotactile information is important to the perception of texture in haptic rendering, and consequently, our system is designed to capture this information.

Multiple projects [4], [15]–[17] have acquired estimates of physical properties using various scanning facilities. However, the cost and expertise required to operate such a facility is a major obstacle in its acceptance. A robotic measurement facility is also inherently inflexible due to workspace limitations, robotic control, and the direct contact between a robot and an object to be scanned. For example, Pai et al. [4] developed a “one stop shop” scanning facility capable of measuring several of the physical properties discussed above. This system consists of an expensive scanning setup and careful planning is required to perform measurement. The facility is not mobile and has a constrained workspace. A “human-in-the-loop” system can be far more flexible because of the superior control a human can exert on a hand-held probe. A human can perform in-the-loop acquisition planning and direct the data acquisition to relevant detail. Additionally, a human-held scanning device can be made quite inexpensively at a small fraction of the price of a professional 3D scanner. Recently, quite a few human operator based systems have been proposed for the acquisition of geometry and visual attributes (e.g., [18], [19]). Systems have also been developed to scan the visco-elastic properties of liquids and semi-solids, deriving estimates for intrinsic physical properties through human interaction [17], as well as to scan quasi-static materials estimating a data-driven deformation model [20]. We believe that interactive scanning of physical object properties can eventually make scanning physical interaction behaviour a routine part of the 3D scanning pipeline.

Many approaches exist to measure surface profiles of samples, and commonly a surface profilometer is employed. Costa and Cutkosky [21] used an optical profilometer to scan rock surfaces, and synthesized similar textures using a fractal algorithm. Wall and Harwin [22] used a linear variable differential transformer to measure displacement of a probe as it is moved (at near constant speed) over a surface. This system generates plausible profiles, however a fixed assembly is required (linear track, motors) and scanning is limited to planar surfaces. One of the disadvantages of using the setup of Wall and Harwin, or a commercial surface profilometer, is the lack of force measurement. Such devices can measure roughness, but not the hardness of a surface. The WHaT [23]
tactile probe in our work measures force and acceleration and we show how these measurements can be used to estimate roughness and stiffness. Other approaches have measured vibrotactile data during manual interaction by attaching accelerometers to robotic linkages [15]. Surface texture could also be measured with an array of pressure sensors. This kind of sensor has been used for tactile sensing in robotics and haptics [24] and the resolution of recent experimental devices reportedly approach the resolution of a human fingertip [25].

The modeling of surface texture is also an area which is receiving increasing attention. These rendering approaches can be coarsely classified as correlated or uncorrelated stochastic, basis function, image-based, and approaches based on digital half-toning [26]. Haptic rendering algorithms for force-based displays often separate the forces into individual components: forces due to interaction with texture, forces due to interaction with the object boundary, and frictional forces. A unified rendering equation is used to compose the individual forces into individual components: forces due to interaction with texture, forces due to interaction with the object boundary, and frictional forces. A unified rendering equation is used to compose the individual forces into individual components. We follow a similar approach.

Wall and Harwin [22] developed a procedural model which incorporates the spatial frequency and amplitude information of a surface, obtained by the discrete Fourier Transform of an object’s surface profile. Okamura et al. [15] reproduce texture detail using basis functions generated based on real-world interaction. In their work, the acceleration profile for a probe is captured as it is dragged over a textured surface, which is then used to determine parameters for a basis function (in this case, an exponentially decaying sinusoid). Textures were synthesized by replaying the basis functions as vibrotactile coding to users via a haptic interface. Basdogan et al. [27] use an image-based approach for haptic texturing, wherein a surface height field is calculated from grayscale image data using static, procedural, and fractal synthesis techniques. However, unlike Minsky’s image-based rendering [28], Basdogan’s rendering algorithm combines the image gradient to perturb the surface normals of the textured object being displayed. This approach is based on a graphical rendering technique called bump mapping, introduced by Blinn [29].

A technique which departs from point-based haptic texture rendering is the work of Otaduy and Lin [30], [31]. They compute texture forces using probe geometry and surface detail. In their experiments, Otaduy and Lin allowed users to explore textured surfaces using a virtual probe which had been textured using depth images. The perceived texture could change drastically depending on the size, shape, and texture of the probe. However, it is also computationally challenging, especially for complex geometry. Shopf and Olano [32] have created a haptic shading framework which builds on these techniques and allows user-defined procedures to define texture characteristics of an object. Their goal was to provide a shading framework for 3D objects similar to graphical shaders, such as the Renderman Shading Language [33]. Our system remains compatible with the techniques discussed in this section in that we synthesize both a texture and compliance estimate.

2 System Overview

Figure 2 contains a data flow diagram showing all the components of our system. At the core of our haptic texturing system are a tactile probe and a visual tracker that stream force-acceleration and position-orientation profiles, respectively, to a host PC. Sensor data is processed offline using a number of steps as shown in the diagram. The processed sensor data are used to generate texture detail for the scanned object, represented as surface profiles. To scan an object, the operator places the object in front of the digital camera and, while the tip of the probe remains in contact, scans the desired section of the object’s surface. Figure 1(a) shows a wooden cobra statue being scanned. If the scan is unsatisfactory (e.g., heavy sensor noise), the section of the surface may simply be scanned again. This is one of the great advantages of a human-in-the-loop system. The operator is able to identify regions of interest of an object’s surface and scan them accordingly.

Figure 1(b) shows a screen shot of the application used to acquire texture detail. On the upper-half of the window, a real-time video obtained from the visual tracker is overlayed with the position and orientation of the probe tip. On the lower-half, the real-time acceleration and force profiles from the WHaT are shown together. Position, force, and acceleration data are streamed to the host PC for offline processing in a later step. Although online generation of textures would be desirable, we found the raw acceleration and force profiles to be sufficient feedback to the operator, e.g., it is easy to notice if the force sensor becomes saturated or aliasing of the acceleration peaks occurs.

Optionally, a geometric 3D model (obtained using an external scanning system) of the scanned object may be used to register synthesized textures to a global frame. Rather than utilizing a high-resolution 3D scanner, we demonstrate that our system can operate with low-resolution image-based models. Various open-source and services free of charge are available for the creation of image-based models (e.g., ARC 3D offered by the EU EPOCH project [35]). The results shown in this paper (see Figure 2) are obtained with an off-the-shelf stereo vision system (Point Grey Research Bumblebee), but we expect similar results with monocular images in a structure-from-motion approach.

2.1 Tactile Probe

We employ the WHaT [23] as a tactile probe (see Figure 1(a)) because it is a mobile and inexpensive device. In our system, the WHaT senses small-scale motion.

1. Preliminary versions of our system have been previously described [6], [34].
Visual Tracking

Fig. 2. The haptic texturing pipeline.

Texture Profile Generation

Stiffness Estimation

Fig. 3. Coarse mesh acquisition based on stereo input. The black-and-white stereo images are displayed as red and green channel in Fig. 3(a). Figure 3(c) shows the corresponding 3D pointcloud and Figure 3(d) the reconstructed and simplified mesh based on 16 stereo pairs. Meshes are processed in Geomagic Studio and MeshLab.

and force data as it is moved across the surface of an object. The probe consists of a force sensor, aligned with the major axis of the device, and a pair of bi-axial accelerometers which provide motion data in 3D space. Each sensor uses 8-bits to represent its data, giving a resolution of approximately 0.015g (where g is the gravitational acceleration) for accelerometers and 0.019N for the force sensor. Force and acceleration sensor data is streamed over a wireless link at a rate of 500Hz to a host PC. No online processing of the WHaT data is performed, but offline processing occurs at the texture profile generation stage (see Section 2.3).

2.2 Visual Tracking

We employ ARToolkitPlus [36], a C++ library for tracking fiducial markers with images obtained from a digital camera. A single stationary camera (Point Grey Dragonfly Express) is used to estimate the position and orientation of the touch probe during the scanning procedure. Figure 1(a) shows the WHaT with attached markers during a scan. From our trials, the visual tracker component is able to estimate the global position and orientation of the touch probe at a rate of approximately 100Hz using this setup. This frame rate is limited by the speed of the tracking software.

The problem of estimating the pose of a calibrated camera (or equivalently a known target) is referred to as the Perspective-n-Point (PnP) problem in computer vision. ARToolkitPlus solves the PnP problem with the robust planar pose (RPP) method of Schweighofer and Pinz [38]. RPP in turn utilizes the iterative method of Lu et al. [39] which is regarded as the gold standard in solving the PnP problem. However, we found the

2. We have employed ARTag [37] in our earlier work but switched due to licensing issues. Our new off-line approach achieves more than double the frame-rate and is more accurate.
TABLE 1
Tracker accuracy comparison with RPP and continuous initialization of iterative solution method of Lu et al. [39] and different number of points. (5 points uses centers of fiducial markers only while 25 points uses centers plus corners. ARToolkitPlus for comparison only since a different data set is required.)

<table>
<thead>
<tr>
<th>Initialization</th>
<th># Points</th>
<th>$Var_x [\text{mm}]$</th>
<th>$Var_y [\text{mm}]$</th>
<th>$Var_z [\text{mm}]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARToolkitPlus</td>
<td>5 targets</td>
<td>$3.51e+00$</td>
<td>$4.38e-01$</td>
<td>$3.42e+00$</td>
</tr>
<tr>
<td>RPP</td>
<td>5</td>
<td>$2.78e-01$</td>
<td>$3.09e-02$</td>
<td>$9.36e-01$</td>
</tr>
<tr>
<td>Continuous</td>
<td>5</td>
<td>$2.52e-01$</td>
<td>$3.19e-02$</td>
<td>$9.75e-01$</td>
</tr>
<tr>
<td>Continuous</td>
<td>25</td>
<td>$8.21e-02$</td>
<td>$1.43e-02$</td>
<td>$3.45e-01$</td>
</tr>
</tbody>
</table>

accuracy of the position estimated by ARToolKitPlus unsatisfactory for our application mainly due to the low number of target points and tracking speed. Instead of using ARToolKitPlus to solve the PnP problem on-line, we only detect marker positions in the images on-line and store them for later processing. As a result our frame rate increased from about 30fps to about 100fps.

We solve the PnP problem by using the iterative method of Lu et al. with the centers and corners of five fiducial markers. Because of the high frame rate, we can simply use the camera pose in the prior frame as an initial guess. We only need to rely on RPP to find a robust starting point either at the beginning of a scan, or if the pose in the prior frame results in a large error in the current frame, indicating the target has moved rapidly. We evaluated our approach by tracking the probe as the user followed the circular indentation on top of a beverage tin can multiple times. We repeat the scan three times at distances of about 40cm, 45cm and 50cm away from the camera, respectively. We fit a circle to the position measurements and calculate the co-variance matrix. Table 1 lists the results for different variants of our approach confirming that using a higher number of points and continuous estimation greatly decreases the variance.

We use simple outlier rejection and a linear Kalman filter [40] to improve the estimate of the position and orientation of the probe further. The outlier rejection compares a position with its neighborhood weighted by a Gaussian. If the difference is larger than a threshold, the position is discarded and the interpolated position used instead. Figure 4 shows the result of tracking the tin can at about 45cm from the camera with the calibration setup. For the circle, we measured a radius of $23.29mm$ with calipers, the tracker estimated a radius of $23.30mm$.

The co-variance matrix of the visual tracker is

$$ Cov = \begin{bmatrix} 8.21e-02 & 5.07e-03 & 1.25e-02 \\ 5.07e-03 & 1.43e-02 & -2.94e-02 \\ 1.25e-02 & -2.94e-02 & 3.45e-01 \end{bmatrix} $$

with all dimensions in millimeters. As expected the main error of the tracker is in the z-coordinate which is the depth. We can also conclude that the tracker should be used to acquire profiles fronto-parallel to the camera and not by scanning away from or towards the camera.

Fig. 4. Tracker position after outlier filtering and Kalman smoothing. A planar circular path is traced multiple times by the probe. Data is shown in camera coordinates with the nodal position of the camera at $(0, 0, 0)$.

A last optional step in the visual tracking pipeline (see Figure 2) is the registration of the scanning trajectory with the surface geometry of an object being scanned. Registration transforms the scanning trajectory from one or more scans into a single coordinate system (e.g. the coordinate frame used by a polygonal mesh). We obtain our coarse geometry by an image-based method without expensive 3D scanners. The details from the scanning trajectory enable us to texture the coarse mesh with haptic detail rivaling high resolution scanners (see Section 4.1 for a comparison). We utilize the iterative closest point (ICP) algorithm [41] implemented in MeshLab [42] for cross-registration. The ICP algorithm as implemented in MeshLab requires two meshes for registration. We turn the scanning trajectory into a mesh by repeating the scanning trajectory with a small offset in the plane normal to the z-axis of the WHaT and triangulating it. This assumes that the user holds the WHaT approximately normal to the surface which is a valid assumption in our set-up. It is well known that the ICP algorithm requires a
good initial estimate to converge to the correct solution. We currently specify the initial transform manually (see Figure 5 for registered scans).

2.3 Texture Profile Generation

In this section, we discuss the post-processing steps employed by our system to arrive at a texture profile from sensor data. Sensor data is collected from the WHaT and the visual tracker at uncoordinated, non-uniform sampling intervals. We synchronize the data from both sensors employing a first-order linear interpolation to re-sample the data at a common period of 1 ms. The orientation of the probe, retrieved from the tracker, is stored as a quaternion and spherical linear interpolation (SLERP) [43] is used to determine intermediary values. Synchronized data is essential when fusing the information from the two sensors for the final profile generation. Before this step outliers are removed, which are caused by communication errors between the WHaT and the host computer and errors in the tracking.

Since the operator is free to use the probe in arbitrary orientations, the pose of the probe relative to the scanned surface must be recovered in order to determine the acceleration of the probe tip, \( \vec{a}_{n,i} \), in a direction normal to the surface at each sampling step, \( i \). The surface profile is generated based on the probe’s motion in a normal direction relative to the surface being scanned. The normal vector, \( \vec{n}_i \), is obtained by registering the path of the probe tip with a 3D mesh. If object geometry is unavailable, the probe is assumed to be oriented orthogonal to the scanning surface and tracker orientation values are used to estimate the surface normal. After registration we remove the gravitational acceleration because then we know the global orientation of the probe. Given the normal vectors along the probe path and the orientation of the device along with its sensor coordinate system, the acceleration is projected into the direction of the unit normal in the camera or world coordinate frame. The normal, the coordinate transform, and the acceleration will change in general at each point \( i \) on the filtered path

\[
\hat{a}_{n,i} = \left[ C \hat{n}_i \cdot \left( \frac{C R_t S \hat{n}_i}{C R_t S} \right) \right] C \hat{n}_i.
\]

We have experimentally observed that the measured acceleration depends on the user applied force – if the user presses the probe firmly against the surface, the acceleration profile, and correspondingly the resulting texture, will contain amplified features; if the user presses lightly, surface features are attenuated. Because it is nearly impossible for a human operator to apply a constant force as they scan the surface, variations in the force profile occur frequently. To adjust for this phenomenon, which we simply name force variation, we apply a scaling factor to the acceleration profile. Its value is based on the observed linear relationship between force, \( f \), and RMS acceleration, \( \text{rms}_{\text{rms}} \), \( \text{rms}_{\text{rms}} = k_c \cdot f + c \). An experimental justification of the linear relationship can be found in Section 3. The scaling factor then becomes

\[
s_i = \left( \frac{\bar{f} + c/k_c}{f_{\text{smooth}} + c/k_c} \right),
\]

where \( \bar{f} \) is the mean force and \( f_{\text{smooth}} \) is a low-pass filtered version of the force profile. The adjusted acceleration value, \( \hat{a}_{n,i} \), is calculated as \( \hat{a}_{n,i} = s_i \hat{a}_{n,i} \).

At this point, we begin generating the height profile by using a Verlet integration scheme [44]. The filtered acceleration values, \( \hat{a}_{n,i} \), are integrated over the interpolated time vector, \( \Delta t \), such that

\[
h_i + 1 = 2h_i - h_{i-1} + \left| s_i \hat{a}_{n,i} \right| \Delta t \\
\Delta t = (t_i - t_{i-1})
\]

where \( h \) is the resulting height profile. It is well known that double integration of acceleration data is prone to drift due to quantization and measurement errors. We use a high-pass (fifth order) Butterworth filter with a cut-off frequency of 1.0 Hz to remove the low-frequency drift from the acceleration measurements before integration.

The height estimate is mapped into the spatial domain using low-frequency components extracted from the scanning trajectory. Since the scanning trajectory and the acceleration data are synchronized in the time domain, the correspondence for mapping is easily found. However, the scanning trajectory is a 3D curve, whereas the height profile corresponding to texture is a scalar function. We find the average surface by applying a low-pass filter to all components of the scanning trajectory vectors. The resulting profile is decomposed into motion normal to the surface of the object, and motion in the tangent plane of the surface. The distances in the tangent plane are used to map the height estimate from the time domain to the spatial domain.

We have now two estimates of the surface height profile: one based on the accelerometer data from the WHaT, and one from the visual tracker. These profiles have different properties – the scanning trajectory from the tracker has good absolute error with high-frequency noise while the trajectory from the accelerometers has low-frequency drift but good high frequency resolution. We fuse the profiles by adding the height estimates after low-pass filtering the trajectory profile and high-pass filtering the height estimate from the WHaT. We choose the transition frequency for the profiles by inspecting the power spectrum. We pick the transition frequency such that the low-frequency content is removed from the WHaT profile but content above what is present in the trajectory profile is maintained (See Figure 12 for examples of the Lomb [45] periodogram).

Next, we describe our approach to synthesize textures using the height profiles.

2.4 Texturing

Creating texture coordinates for the height profile involves two steps: (i) registering the scanning trajectory with a 3D geometry, and (ii) generating texture coordinates on the surface of the 3D object. We have reported
our method in detail in [7] and only briefly summarize it here for completeness.

Fig. 5. Scan registration for the cobra statue. Shown is the coarse geometry of the image-based reconstruction.

Texture coordinates are assigned to elements of a 3D mesh directly from the set of mesh points resulting from scan path registration by the ICP algorithm (see Figure 5 for our example of the cobra statue). Two dimensional texture coordinates are generated for each vertex of a polygon. We consider the signed distance along the scan path to be one of the parameters of the texture function. The other parameter we estimate as the signed minimum distance to the scan path, (e.g. between mesh vertices and scanning segments) with positive and negative direction being arbitrarily, but consistently, chosen.

3 Compliant Surfaces

Compliance represents the force-displacement relationship of a deformable surface and is the inverse of stiffness. Measurement of the force-displacement relationship has been conducted in the context of haptics before [4], [15], [20], [46]. In these approaches (cf. Okamura), the surface’s compliance value may be computed by deforming a surface (e.g., by pushing against it) and directly measuring the force generated at the contact point and the resulting displacement. It should be noted that while it is theoretically possible to directly measure this force-displacement relationship based on sensor data from the WHaT, this is not possible in practice.

Direct measurement of the force-displacement relationship with the WHaT is not a practical option for two main reasons: the inability to estimate precise surface displacement for slow motions, and measuring large forces with a hand-held device. The estimation of displacement based on accelerometers is limited due to drift as previously discussed. These errors are aggravated in measuring slow motions as would be necessary in order to perform this type of quasi-static force-displacement measurement. Furthermore, many solids for which we acquire deformable haptic texture are quite rigid (e.g., cork, pencil eraser) and would require large forces to deform beyond the range possible with the WHaT. In addition to the limitations of the WHaT sensors, the visual tracker is not capable of measuring displacement of the probe tip with a degree of accuracy for which a reasonable force-displacement relationship can be determined. Instead, we develop an approach based on the dynamic behavior during surface scanning.

3.1 Dependency of Acceleration on Force

Before introducing our approach of estimating a measure of surface compliance, we will present our observations about the experimental relationship between the acceleration and force measured by the WHaT sensor. As briefly mentioned in Section 2.3, we observe a proportional increase in acceleration amplitude with the force applied by the operator. Figure 6 shows the force and acceleration profiles while a user repeatably scans the surface of a stiff pencil eraser and changes the applied force. Clearly, higher forces result in higher accelerations. This is not surprising considering a second-order linear dynamic system

\[
m\ddot{u} + b\dot{u} + ku = f_{ext}
\]

with stiffness \(k\), damping \(b\) and displacement \(u\) and external force \(f_{ext}\). This can guide our understanding of the observations, but it is unlikely to be an accurate description of the contact behavior. It is also of no help in estimating material parameters (e.g., \(b\) and \(k\)) since we neither observe \(f_{ext}\) nor \(u\).

In order to clarify the relationship between the force \(f\) and the acceleration \(a\) measured by the sensor, we plot the root-mean-squared (RMS) acceleration over the low-pass filtered force. We opt to low-pass filter the force in order to obtain a force which may be considered proportional to the average applied force by the user during a scan or portion of a scan. In our experiments, we used a first-order Butterworth filter with a cutoff frequency of 2.5Hz. The RMS acceleration \(a_{RMS}\) is

\[
a_{RMS} = \sqrt{\frac{1}{N} \sum_{i} a_i^2}
\]
of displacement are controlled by the damping ratio \( \zeta = \frac{b}{2\sqrt{km}} \). The damping ratio is inversely dependent on the square-root of the stiffness. Based on this argument, we expect to see smaller accelerations with the increased damping ratio of more compliant surfaces.

We quantify the effect by the same line fitting procedure described in Section 3.1. We expect the linear coefficient to be proportional to the stiffness of the material, or equivalently, we expect that locally, we can approximate the dependency of the acceleration on the force linearly. However, we note that the standard deviation of fit for the line in the case of sandpaper is high and that scanning with an increasing amount of force is difficult for many surfaces. Therefore, we propose a scanning procedure for compliance which consists of repeatably scanning the same surface with different amounts of applied force. The advantage of this scanning procedure is the larger number of samples in general and the easier variation of applied force. Raw sensor data for such a repeated surface scan is shown in Figure 6. Note, as well, that the profile again clearly shows the dependency of the acceleration on the applied force. We devised a simple segmentation routine which cuts out sections of the data where the force sensor does not record any force or the acceleration is constant suggesting that the probe is stationary. The segmented profile is shown in Figure 9.

Fig. 7. Force and acceleration during scanning of sandpaper. Top plot shows the measured force (grey, light) and low-pass filtered force (brown, bold) used for surface compliance estimation. The user applied force increases with time (from left to right).

Fig. 8. Linear approximation of the dependency of acceleration on force. Error bars indicate the inverse weight of a sample in linear least squares. Nearly rigid sandpaper in Figure 8(a), compliant pencil eraser in Figure 8(b).

Fig. 9. Segmentation result for repeated scans of an eraser; Figure 6 shows the input data.

The described scanning procedure for the eraser results in the line fit shown in Figure 8(b) and a linear coefficient \( k_c = 0.1793 \) with a standard deviation of \( \sigma = 0.1121 \). These values support our argument; the linear coefficient is smaller than for stiff sandpaper and the scanning procedure of repeat scans results in a lower standard deviation than one scan with varied force. However, we note that our approach depends on the occurrence of acceleration bursts caused by surface roughness, i.e., completely smooth surfaces will exhibit only very small accelerations no matter what force is applied. As a result, very smooth surfaces will not be usable in our estimation procedure of compliance. We

### 3.2 Compliance Estimation

The above observations about the relationship between dynamic force and acceleration helps us to find a measure of surface compliance. We start by noting that tapping on a hard surface causes vibrations while tapping on soft materials does not. It is therefore reasonable to expect that hard materials cause high deceleration values as the WHaT probe impacts them, while soft materials cause lower values. In the second order system of Equation 1, the vibrations based on sudden variation
present more complete results in Section 4.2 including a comparison to independently measured material stiffness and Young’s moduli.

4 Results

In this section we present the results of using our system to scan real-world objects. In addition to showing the haptic textures we generate, a comparison is done using scans from a high-resolution laser scanner. We refer the reader to our earlier work [7] for additional haptic textures, including scanning stochastic surfaces such as sandpaper and cork. Compliance estimates are also obtained for a variety of materials, perceptually ranging from hard to soft. These results are obtained with our new method for estimating the stiffness coefficient and we compare the results to measurements of stiffness and the object’s reduced Young modulus from an independent measurement setup (described in the Appendix).

4.1 Surface Profiles

![Fig. 10. High-resolution 3D mesh of the cobra (shown in flat shading).](image)

Here, we present the results of scanning different areas of the wooden cobra statue shown in Figure 2. Our method does not require a 3D scanner but only a coarse 3D model which can either be obtained with image-based methods as presented in this paper or manually. We employ high-resolution scans in the following only for comparison and clarification. The statue has three distinctly textured areas: on the back of its hood, the front of its body and the front of its hood. The textures can be visually identified in the meshes obtained with high-resolution laser scans shown in Figure 10. The laser scans are obtained with a Minolta Vivid Vi-910 scanner with a tele-lens for which Minolta specifies an accuracy of $\delta x = \pm 0.22\text{mm}$, $\delta y = \pm 0.16\text{mm}$ and $\delta z = \pm 0.10\text{mm}$ [47]. Scans were taken with the object on a turntable at $30^\circ$ steps and have been merged with Geomagic Studio, a commercial software packaged with the scanner.

In order to compare the profiles with the 3D scans, we register the profiles to both, the coarse 3D model and the 3D high resolution mesh with the approach described in Section 2.2. We process the sensor data from our system as discussed in Section 2.3 using the normals from the coarse 3D model and determine the direction of the gravitational acceleration by the global orientation of the coarse 3D model. The registration with the high-resolution mesh is used only to be able to compare the height profile with the high-resolution mesh. Figure 11 shows the comparison for the three selected areas of the statue. Our system performs satisfactory in all three comparisons since the characteristic features of the surfaces are visible and the feature locations between the high-resolution 3D mesh and the profiles obtained with our system align. In comparison to our earlier work [6], there is a drastic improvement in the localization of the features. This improvement has been the result of the new sensor fusion approach utilized in our current system which in turn has been possible due to the improved visual tracking capabilities.

Our approach to sensor fusion is quite simple, we determine a spatial transition frequency between the height-profile obtained with the tracker and with the WHaT. The motivation for this transition is that the height-profile from the tracker has good absolute accuracy but the WHaT has superior higher frequency relative accuracy. The lower frequencies from the WHaT are polluted by drift. In Figure 12, we show the spatial power spectra for three scans of the wooden cobra in Figure 11. The power spectrum of the tracker is the complete spectrum while the WHaT spectrum is only the part above the transition frequency which enables a direct comparison of the power in higher frequencies between the two profiles. For all three profiles, the WHaT spectrum adds higher frequencies to the fused profile. However, the spectrum plots can not reveal if these higher frequencies are simply noise or part of the profile. Figure 11 shows the fused profiles in comparison to the tracker-only profiles, and it reveals that the WHaT helps to flatten the top of the peaks in the profiles and bring the profiles closer to the 3D mesh profile. We believe the visible error to be caused by two facts: On the one hand, our system will produce large acceleration spikes during traversal of sharp features in a profile and the probe may even lift-off from the surface for a short period of time. On the other hand, there may be some amount of synchronization error between the WHaT and the tracker. We believe the synchronization error to be caused by high frequency variations in the scanning velocity which are not captured by the tracker.

4.2 Compliant Surfaces

In the following, we will present the results of our dynamic stiffness estimation procedure for a set of objects with different stiffness ranging from a soft mouse-pad to sandpaper. We start our evaluation by measuring baseline values for the stiffness of the materials in our test set. We conduct a compression test by probing each object (except the sandpaper) with a spherical indenter with a radius of $R = 6.265\text{mm}$. The indentation is executed with
(b) Front hood profile.  

(c) Back hood profile.  

Fig. 11. Height profiles for the wooden cobra. The height profile from our system are shown with solid (purple) lines, top row is from the tracker only, bottom row are fused profiles from visual tracking and the WHaT. The dashed (gray) lines are the height profile obtained based on the high-resolution laser scans for comparison. A height profile is the deviation from the average surface.

![Graphs showing height profiles for different parts of the cobra.](image)

(a) Front body profile.  

(b) Front hood profile.  

(c) Back hood profile.  

Fig. 12. The power spectral density of profiles in Figure 11. Dashed (gray) and solid (purple) lines on the left are the spectra of the height profiles from the tracker and from the WHaT, above the transition frequency of the fusion algorithm, respectively. The spectra of the fused profiles are shown on the right. The spatial transition frequencies are on average $0.59 \frac{1}{\text{mm}}$, $0.41 \frac{1}{\text{mm}}$ and $0.86 \frac{1}{\text{mm}}$, respectively. The spectra are calculated by the Lomb [45] periodogram.

![Graphs showing power spectral density for different parts of the cobra.](image)

TABLE 2  
Stiffness estimates $k_c$ and standard deviation of fit $\sigma$ for several materials. Experimental verification from a spherical indentation test. Proportionality factor $\tilde{k}_{\text{comp}} = \frac{f}{u^{3/2}}$ and reduced Young’s modulus $E_r$.

<table>
<thead>
<tr>
<th>Material</th>
<th>$k_c$</th>
<th>$\sigma$</th>
<th>$\tilde{k}_{\text{comp}}$</th>
<th>$E_r$ [MPa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>mouse pad</td>
<td>0.108</td>
<td>0.020</td>
<td>0.276</td>
<td>0.083</td>
</tr>
<tr>
<td>foam B</td>
<td>0.108</td>
<td>0.059</td>
<td>0.212</td>
<td>0.063</td>
</tr>
<tr>
<td>foam A</td>
<td>0.115</td>
<td>0.310</td>
<td>1.83</td>
<td>0.548</td>
</tr>
<tr>
<td>sponge B</td>
<td>0.117</td>
<td>0.055</td>
<td>1.56</td>
<td>0.469</td>
</tr>
<tr>
<td>sponge A</td>
<td>0.132</td>
<td>0.054</td>
<td>3.29</td>
<td>0.986</td>
</tr>
<tr>
<td>toy (foam)</td>
<td>0.157</td>
<td>0.102</td>
<td>4.02</td>
<td>1.21</td>
</tr>
<tr>
<td>floor felt</td>
<td>0.163</td>
<td>0.038</td>
<td>9.14</td>
<td>2.74</td>
</tr>
<tr>
<td>pencil eraser</td>
<td>0.179</td>
<td>0.112</td>
<td>13.2</td>
<td>3.97</td>
</tr>
<tr>
<td>toy (crocodile)</td>
<td>0.266</td>
<td>0.145</td>
<td>25.8</td>
<td>7.74</td>
</tr>
<tr>
<td>rubber tire</td>
<td>0.526</td>
<td>0.096</td>
<td>39.4</td>
<td>11.8</td>
</tr>
<tr>
<td>cork</td>
<td>0.853</td>
<td>0.257</td>
<td>49.0</td>
<td>14.7</td>
</tr>
<tr>
<td>sandpaper</td>
<td>0.907</td>
<td>0.725</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

TABLE 3  
Material properties for various objects. A probe consisting of a 6D force-torque sensor (ATI Nano 25) mounted in a haptic device (MTB Freedom 6S). We compress the material by manually applying force and measuring the force with the force-torque sensor and the position with the haptic device. See the Appendix for a more detailed description of the set-up. We repeat this test a number of times and estimate a stiffness and the reduced Young’s modulus of the material. We do not employ the well-known method by Oliver and Pharr [48] for the estimation since their method assumes plastic deformation of the surface during the indentation test. Our materials are soft and the forces small and we can assume elastic deformation. We employ a full non-linear fit to the loading curve in a similar fashion as proposed by Briscoe et al. [49]. Both methods are based on the Hertzian contact model for a sphere-plane contact given by:

$$f = \frac{4}{3} \sqrt{R} \frac{E}{1 - \sigma^2} u^{\frac{3}{2}} = \frac{4}{3} \sqrt{R} E_r u^{\frac{3}{2}} = \tilde{k}_{\text{comp}} u^{\frac{3}{2}}$$  \hspace{1cm} (2)$$

$E$ and $E_r$ are the Young’s modulus and the reduced Young’s modulus, respectively, and $\sigma$ is the Poisson’s ratio of the object. We assume that our indenter is infinitely stiff as compared to the materials.

$E_c$ and $\sigma$ are the Young’s modulus and the reduced Young’s modulus, respectively, and $\sigma$ is the Poisson’s ratio of the object. We assume that our indenter is infinitely stiff as compared to the materials.
The table shows that our dynamic stiffness estimation procedure has successfully ranked the objects according to stiffness. This enables us to use the stiffness estimates in our haptic rendering application. While we cannot provide an analytic relationship between the stiffness coefficients estimated in our dynamic estimation procedure and the results of the compression tests, an experimental relationship can be easily observed (see Figure 13). We fit an exponential curve with the Matlab routine \textit{fminssearch} taking into account the uncertainty in the stiffness coefficients from our scans. The relationship is

\[ k_c = 0.0217 + 0.0883e^{0.149E_r[\text{MPa}]} \]

There are some limitations to this relationship. The stiffness estimation method for \( k_c \) from our scans assumes that the object’s surface is not smooth since we require accelerations normal to the average surface. The object must also be reasonably homogeneous, e.g., the estimate for packaging foam A is low likely due to the foam having large variations in stiffness due to an internal coarse cell structure. Our scanning procedure signals this issue with a high fitting error. However, we can conclude that our novel technique for stiffness estimation generates a good heuristic regarding the stiffness of scanned surfaces and we were able to establish an experimental relationship between the Young’s modulus of an object and the stiffness coefficient \( k_c \).

![Fig. 13. Relationship between (reduced) Young’s Modulus and experimental stiffness coefficient \( k_c \).](image)

The stiffness coefficients are used to modify the stiffness of textures when they are displayed to users in a haptic interface to a hapto-visual application. We feel that the textures are pronounced and salient texture features are easily detectable if using a very rigid stiffness value, such as the sandpaper. For lower stiffness values, texture features are attenuated, with the soft mouse-pad being mostly characterized by rigid body contact forces.

4.3 Haptic Rendering

We demonstrate our compatibility with the standard rendering algorithms by incorporating several of the techniques discussed in Section 1.1. To display the textured objects, a hapto-visual application that renders textured models to the PHANTOM\textsuperscript{TM} series of devices is used, particularly the Premium 1.0 device since it enables one to explore the haptic textures with higher fidelity. Users can interact with textured 3D objects using a virtual proxy, which is a single point of contact interface metaphor for the haptic device.

We adopt a force model similar to that used by Siira and Pai [50], whereby the overall rendered force is decomposed into those acting along the normal, \( \vec{n} \), and tangential, \( \vec{t} \), directions to a point on the surface of the 3D object where the virtual haptic proxy makes contact

\[ \vec{f} = \vec{f}_c + \vec{f}_t + \vec{f}_\tau = [k_x \Delta x + k_c \Delta h] \vec{n} + [u_s |\vec{f}_c + \vec{f}_t|] \vec{t}. \]

The forces due to rigid body constraints, \( \vec{f}_c \), and texture forces, \( \vec{f}_t \), are rendered using a penalty-based method, whereby spring forces are generated proportional to the penetration depth of the proxy into the surface, \( \Delta x \), or the texture, \( \Delta h \). The frictional force, \( \vec{f}_\tau \), uses a preselected static friction coefficient, \( u_s \), and its magnitude is proportional to the rigid body and texture forces.

Figures 1(c) and 1(d) show a hapto-visual application which is displaying the textured 3D cobra object. Sections of the mesh which have been textured are highlighted, and, when the user brings the proxy in contact with these areas of the mesh, the texture is displayed according to our haptic rendering algorithm.

5 CONCLUSION

Our system is a viable solution for creating textures for use in 3D hapto-visual applications demonstrated by the results of scanning a wooden cobra statue. Our textures are based on surface height profiles fused from data from the visual tracker and from the WHaT probe. We compared the height profiles from our system with profiles calculated from high-resolution scan data. Some spatial warping and scaling are present in our constructed surface profiles but the overall characteristics of the scanned surfaces are well captured. We also have presented an independent analysis of the accuracy of the visual tracking in form of a covariance matrix of the position estimates of the tracker. The most novel aspects of our approach is the estimation of a stiffness coefficient for haptic texturing. Our method estimates this coefficient solely based on the analysis of the force and acceleration during a surface scan and we are able to show an experimental relationship to the contact stiffness as measured during compression.

APPENDIX

The compression test described in the following are employed in the evaluation of our dynamic stiffness
estimation method. The tests are only used in evaluating system performance and are not part of our scanning system. Figure 14(a) shows the setup for the measurement. The force-displacement curves recorded by the set-up in testing the black packaging foam is shown in Figure 14(b). The curves are recorded during repeated slow compressions of the object. The curve shows some non-elastic deformation of the packaging foam which needs a larger recovery time to return to its original shape. We correct the offset value in the subsequent processing. We segment the curves into individual compression curves (see Figure 15(a)) and fit each segmented curve separately. We accomplish the curve fitting of the Hertzian contact model with the Matlab routine fminsearch (see Figure 15(b)). Applying Equation 2 to the stiffness estimation leads to the results presented in Table 2.

![Force-Displacement](image1.png)

![Segmented Tests](image2.png)

![Model fit](image3.png)

Fig. 14. Compression Test. The test is performed to obtain an independent verification of our dynamic stiffness estimation method based on scan profiles. Measurement data in Figure 14(b); set-up in Figure 14(a).

![Model fit](image4.png)

![Segmented Tests](image5.png)

Fig. 15. Fitting a Hertzian contact model. Figure 15(a) shows the data, while Figure 15(b) shows the fit to the Hertzian contact model (purple solid line).

**ACKNOWLEDGEMENT**

The authors gratefully acknowledge the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC), from Point Grey Research and from the Ontario Research Network for Electronic Commerce (ORNEC). The authors would like to thank Vijaya Lakshmi Guruswamy for her help with the experiments, data acquisition and the supplemental video. They would also like to thank Jilin Zhou for his role in the preparation of the compression tests. They also grateful to Dinesh Pai and Peter Rizun for the WHaT.

**REFERENCES**


