Fast Calibration Of Haptic Texture Synthesis Algorithms

Gianni Campion, *Student Member, IEEE* and Vincent Hayward, *Fellow, IEEE*

Abstract—Calibrating displays can be a time-consuming process. We describe a fast technique for adjusting the subjective experience of roughness produced by different haptic texture synthesis algorithms. Its efficiency is due to the exponential convergence of the "modified binary search method" (MOBS) applied to find points of subjective equivalence between virtual haptic textures synthesized by different algorithms. The method was applied to find the values of the coefficient of friction in a friction-based texture algorithm that yield the same perception of roughness as the normal-force variations of conventional texture synthesis algorithms. Our main result is a table giving the perceptual equivalence between parameters having different physical dimensions. A similar method could be applied to other perceptual dimensions provided that the controlling parameter be monotonically related to a subjective estimate.

Index Terms—Haptic Rendering, Haptic Textures, Passivity, Roughness.

I. INTRODUCTION

THE function of programmable sensory displays is to reproduce those aspects of the ambient physics that can satisfy our perceptual system in the accomplishment of specific tasks. For example, reading text on a computer screen requires a minimum level of contrast [1]; color matching requires fine characterization of the device used to mediate the conversion from a signal representation to its physical realization [2]; and so on. In applications where the signal to be displayed is processed digitally, the algorithms responsible for the synthesis of the signal participate in the result as much as the transducers do. In the graphics and audio domains, methods have been proposed to assess and compare algorithms, see for example [3], [4].

Haptics engineers face similar problems [5]. While it is known that force-feedback devices of different makes can produce different perceptual results [6], to our knowledge, there has been no report of suitable methods to perceptually compare different haptic synthesis algorithms or to link the psychophysics of haptic perception to the properties of the algorithms employed. Among the gamut of rendering approaches [7], the synthesis of textures stands out because the input to an algorithm may be specified in a simple and unambiguous manner, i.e., a surface geometry. The methods for converting a geometry into a force field that a user experiences can nevertheless be widely varied.

This wide variation may be explained by the numerous assumptions which must be made regarding the physics of

G. Campion and V. Hayward are with the Haptics Laboratory, McGill University, Montreal, Quebec, Canada.
E-mail: {champ|hayward}@cim.mcgill.ca

mechanical interaction between a probe and a surface. These assumptions yield different approaches to convert a surface geometry into a force field. The same surface geometry can give rise to different fields according to the assumed geometry of a probe tip, as a first example. Besides geometry, many further assumptions must be made regarding the materials in contact, their compliance, the resulting tribology, the statics and dynamics of the probe, the presence of foreign particles, of lubricants, and so on. Despite these numerous possibilities, texture synthesis is achieved today using simplistic algorithms, but since these algorithms are driven by different parameters, direct comparison is not possible.

Works related to the perceptual quantification of texture synthesis algorithms include those of Ho et al. [8]. They studied whether a two-dimensional force field could be distinguised from a three-dimensional force field when rendering square-wave gratings. They concluded that a square-wave grating rendered as a two-dimensional force field was indistinguishable from the three-dimensional version if the height of the wave or its spatial period was smaller than 1.5 mm. Weisenberger et al. studied human detection performance of orientation for three- and two-dimensional textures experienced through different devices [9]. The results support the notion that the only significant difference was between sinusoidal textures and square waves in three dimensions. The authors also compared viscous-based textures with spring-based textures, but, perhaps due to the limited data available, no significant difference was found. The question of the perceptual properties of specific algorithms was briefly mentioned in [10].

During the course of experimentation, we noticed that small changes in a synthesis algorithm could yield tangible differences. This observation raised the question of how algorithms could be compared with regard to the perceptual experience that they produce.

The main purpose of this paper is to describe a fast psychometric method able to calibrate a pair of haptic synthesis algorithms with respect to each other from the view point of one particular perceptual dimension. Efficiency is essential since a complete perceptual characterization of texture is complex [11]. Our study confirms that a texture algorithm based on non-geometrical cues produced by modulating the tangential friction force can elicit a perception of roughness equivalent to that given by a geometry-based algorithm. It is then possible to compare them fairly for their intrinsic properties such as implementation complexity or the range of sensations they can synthesize. Once two algorithms are calibrated to elicit an equivalent sensation of roughness, it is possible to compare them fairly for their intrinsic properties,

despite their reliance on a possibly different set of parameters. In particular, it becomes possible to quantify the relative passivity margin that algorithms provide for an equivalent level of perceptual roughness or conversely, the relative roughness experienced for the same passivity margin.

II. APPROACH

As alluded to earlier, the perceptual calibration of a device or of an algorithm can be a time-consuming process. In general, calibration requires one to cyclically adjust parameters until two perceptual outcomes become subjectively equivalent. It is a search process where the objective can be simply defined but where the search space grows combinatorially with the number of parameters involved. A powerful simplifying property of any search problem is to depend on parameters that are monotonically related to the objective. The monotonicity property guarantees the success of any gradient-based search, that is, in spite of the fact that adjusting one parameter may undo the effect of another, progress toward the goal is guaranteed. By analogy, in color displays, it is by far more convenient to adjust chromaticity using "hue, saturation, and value" parameters than using "red, green, and blue" parameters.

A. Haptic Roughness Studies

Arguably, roughness is the dominant property that one may want to match when substituting one algorithm for another—or one device for another—in a virtual reality haptic simulation. Roughness has been the topic of many studies, using both natural and electromechanically-generated stimuli. A current theory proposes a dual coding underlying texture perception [12]. According to this theory, the perception of microtextures (spatial period smaller than 0.2 mm) would be determined by the vibration of the skin moving against a surface and roughness estimation would follow a perceptual model based on the sensitivity of the Pacinian system. For coarser textures, roughness estimation would depend on the spatial characteristic of a texture. Studies using raised dot patterns and gratings showed a strong influence of inter-element spacing, but have produced conflicting results [13], [14]. The U-shaped spacing-to-roughness function found when appreciating texture with a bare finger shows an inversion point for inter-element spacing of 3.5 mm. The shape of this function seems to be preserved when a probe is used [15], but the inversion point shifts according to the tip geometry and exploration speed, denoting a lack of perceptual invariance [16]–[18].

B. The Question of Individual Differences

Roughness is one of the dimensions contributing to the perception of textures. To explore more thoroughly the different dimensions of texture perception, multidimensional scaling analysis is often employed. Recently, using this approach, Hollins and colleagues revealed that texture perception is subject to significant inter-subject variations [19]. Specifically, the perceptual space of some subjects could be best described with the two dimensions of roughness and compliance, while other subjects responded to a third factor, the sticky-slippery dimension. Studies often examined subjective roughness estimation and discrimination using stimuli that had similar geometries. Here, we used stimuli that were fundamentally different, one being based on geometry and the other on frictional properties. It could therefore be expected that inter-subject differences in texture perception would be exacerbated. Roughness plays a key role in texture perception, but little is known regarding the individual variations in roughness perception, since experimental designs are typically not aimed at detecting inter-subject variations. Yet, virtual environments must work equally well for most people, in spite of these variations. Again the analogy with color screens is apt. Today, whether screen displays use LCD technology, phosphorescence, or plasma cells, commercial devices have achieved a fair degree of perceptual equivalence despite great differences in the raw stimulus and individual preferences.

C. Experimental Precautions

To ensure that the experiment was carried out in the best conditions possible, we first analyzed the minimum requirements which a haptic simulation system must meet to be able to generate predictable textural effects [20]. Secondly, we developed a systematic method to characterize the properties of texture synthesis algorithms for their control and mathematical properties [21]. Third, we picked a pair of texture synthesis algorithms selected for their differences in the approaches to converting surface shape to a force field. Fourth, we ascertained that sensation perception scaled monotonically with the value of these parameters using the same apparatus and algorithms as in the present study [22].

D. System Considerations

For calibration to be meaningful, one must ensure that programmed changes in parameters are reflected in what is felt by the user. Earlier, we described six intertwined inequalities which reflect the ability of a system to synthesize and reproduce given textures [20]. These conditions address the possible creation of artifacts arising from quantization, aliasing, and lack of control passivity; they provide a quantitative descriptions of the constraints imposed on the texture rendering parameters by the hardware characteristics of a given haptic display system. One of these inequalities states that, in order to ensure the passivity of the user interaction with a virtual sinusoidal texture, the product of the spatial frequency by the amplitude must be bounded by a constant. This constraint, has immediate implications for roughness calibration. It implies that for a given algorithm, in the course of an experiment not all combinations of parameters are valid. This inequality also raises the prospect that a pair of algorithms could never be calibrated for roughness, or that regions of achievable equivalent roughness are small.

To avoid the artifacts typical of texture synthesis, we used a highly reliable, custom-made hardware platform called Pantograph, which is fully described in reference [23]. This device meets all the required hardware-related requirements. It can display forces of up to 2 N in a two-dimensional workspace of

100×60 mm; and within a flat bandwidth from DC to 400 Hz, see Fig. 1, making this device well described by a rigid body model without structural dynamics. These figures are typical of the mechanical signals arising from the perceptual interaction with surfaces [24]. The torque commands are processed by a lowpass reconstruction filter, so that the commands to the motors match the mechanical bandwidth of the system.
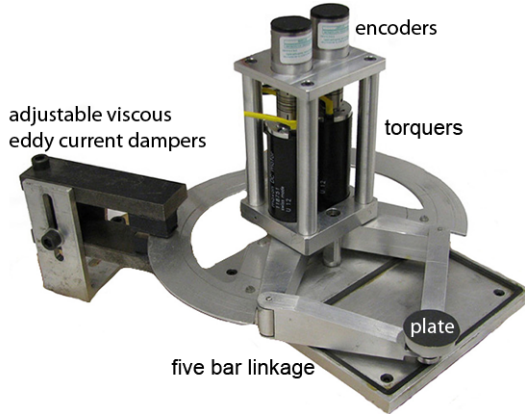

Fig. 1. Apparatus. The Pantograph device has a direct-driven five-bar mechanism that drives a small plate in the horizontal plane. In the present version, each joint was retrofitted with permanent magnet eddy-current viscous damper. The amount of viscosity is adjusted by controlling the amount of overlap between the polar pieces and the rotating annular blades.

With the aim of further reducing possible artifacts and increasing the range of valid parameters, the device was retrofitted with viscous dampers since it is known that adding damping to a system causes it to be more robust and to be less prone to oscillations, see [25] for details. These devices are based on the principle of Foucault current brakes. Two annular blades were concentrically mounted on the proximal arms so as to rotate through the air gap of C-shaped magnetic circuits. In the present configuration, a fixed magnetic field was generated by rare-earth magnets adjusted to give a damping coefficient, B_{a}, of approximately 6 mN·m·s. Interacting with the unpowered Pantograph felt like exploring a viscous field.

To cancel the effects of this viscous field in free space, negative, generative damping was actively produced by the motors to cancel the inherent damping given by the dampers. This technique initially suggested by Colgate and Shenkel [26] was to our knowledge, for the first time successfully applied, thanks to the accurate viscous damping given by eddy current brakes. In each joint, the compensating torque was produced according to

$$\tau_k = +B_{\mathrm{d}}\,\hat{\omega}_k, \tag{1}$$

where the torque τ_k applied at time k was computed from an estimate of the joint angular velocity, $\hat{\omega}_k$, obtained by backward difference over a 24-sample window, and with $B_{\mathrm{d}} < B_{\mathrm{a}}$ adjusted such that the residual viscosity field could barely be felt. These compensating torques were not applied inside the regions where textures were synthesized, thus augmenting the system passivity margin to achieve a virtual wall stiffness of $2\,000$ N·m^{-1}.

E. Psychophysics

In essence, the problem was to determine parameter settings that achieve a *Point of Subjective Equivalence*, or PSE, for a given attribute within a population of subjects. It is well accepted that an accurate method to determine a PSE is to measure complete psychometric functions. In general, a psychometric function relates the value of a single parameter to the probability of a perceptual event to occur, such as the detection of a weak signal or discrimination between different signals. Once such a function is known, the value of the parameter corresponding to the 50% probability determines the point of subjective equivalence. For the task at hand, a yes/no experimental design can be implemented and subjects are presented with a sequence of two textures, one synthesized by a reference algorithm and one synthesized by the algorithm to be calibrated. Subjects then report which of the two virtual surfaces feels rougher. Based on the responses the psychometric function can be identified.

1) Limitations of the Method of Constant Stimuli: A sampling method, called method of constant stimuli, is to design the stimuli presented to the user before the experiment independently from the answers. In general this approach is slow, because hundreds of trials are usually necessary to estimate a single psychometric function. For the present problem of calibrating a sinusoidal texture profile synthesized with two different algorithms, at least two parameters, i.e. amplitude and spatial frequency, need to be varied. A minimal experiment would require to test for at least three values of each parameter. This approach is therefore not viable when the objective is to quickly calibrate a machine, possibly specifically for each individual.

2) Advantages of Staircase Methods for Calibration: Another approach is to employ adaptive staircase methods because they are designed to find specific values of the psychometric function, not the whole function. Instead of uniformly sampling the psychometric function, most of the trials are concentrated in the neighbourhood of a target value by changing the stimuli presented based on the subject's answers. For example, the common 1-up 1-down method, converges to the value corresponding to the 50% probability. Even faster results can be obtained variable step adaptive methods, which further reduce the sampling requirement. Staircase methods (both with fixed and adaptive step size), however, do not allow for a reliable reconstruction of the complete psychometric function because its graph is coarsely sampled [27].

3) Modified Binary Search Method: Among the family of variable step staircase methods, the "MOdified Binary Search" method, or MOBS [28], is designed to converge exponentially to the 50% point of the psychometric function using an adaptive step size. It relies on binary search, while accounting for threshold migration during an experiment. It was originally designed for visual contrast detection and can be adapted to yes/no experimental designs. The main reasons for selecting this method are its speed and simplicity. MOBS has also a characteristic which is highly relevant to haptic texture calibration. The limitations of any haptic device, such as resolution and damping, create tradeoffs between the parameters of the textures that can be reliably synthesized without

artifacts. This problem impacts any thresholding process since a sought threshold can fall outside the range of admissible parameters. MOBS can cope with this occurrence because it keeps tracks of multiple candidate intervals and because its fast convergence allows for repetitions of the same thresholds. Finally, MOBS' performance is at par with more complex and rigorous thresholding approaches [29]. Taken together, these properties have advantages that greatly outweigh those of other methods.

The first step of MOBS is to define an interval that should contain the value of the parameter to be thresholded. The experiment proceeds by testing the midpoint of this interval, which is then reduced by dichotomy based on the answer of the subject. To account for threshold migration, the algorithm tests the boundary of the interval when the subject does not reverse her or his answer twice in a row. If the threshold falls outside the current interval, the method rolls back to a previous boundary. At any time, a history of 3 upper and 3 lower bounds is kept in two stacks; the update rules for the stacks can be found in [28]. The termination criterion is based on a fixed number of reversals as well as on the size of the interval. If a prescribed number of reversals is reached but the interval is too large, the trial is extended for two more reversals; when the trial ends, the estimated threshold is the midpoint of the last interval. For a complete description please further refer to [29].

III. MATERIALS AND METHODS

We used two texture synthesis algorithms that were as different in concept as possible. We then asked subjects to calibrate one algorithm against the other using the approach just described in order to find parameters that produced an equivalent sensation of roughness. In doing so, we never asked subjects to directly estimate roughness with either of the two algorithms. At the conclusion of the experiment it was possible to find such parameters, but other aspects of the texture synthesis differed between the two algorithms.

A. Algorithms

Textured walls were created in two dimensions, x, z. The point $p = [p^x, p^z]^\top$ represents the position of the probe in the virtual space. A geometrical profile $h(p^x)$ was applied to a straight virtual wall of stiffness κ_0. Please refer to Fig. 2 for an illustrative description of the algorithms.

Algorithm **A** produces a force field according to:

$$f_{\mathbf{A}}(p) = \begin{cases} -\kappa_0[0, (p^z - h_{\mathbf{A}}(p^x))]^\top & \text{if } (p^z - h_{\mathbf{A}}(p^x)) < 0, \\ [0, 0]^\top & \text{otherwise,} \end{cases}$$
$$h_{\mathbf{A}}(p^x) = A\sin(2\pi p^x / L). \tag{2}$$

where the penetration, $d = p^z - h_{\mathbf{A}}(p^x)$, is computed from the boundary of the texture along the z direction, A is the amplitude and L is the spatial period. Algorithm **F** is based on the modulation of dry friction to generate a texture. A time-free friction model, described in [30], is used to generate
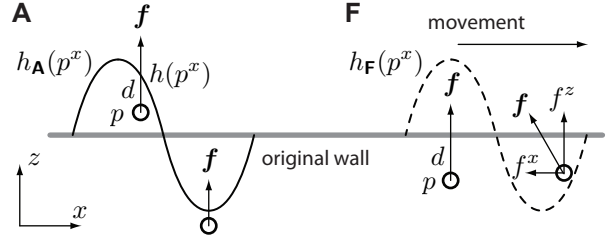

Fig. 2. Algorithm **A** produces a force field aligned with the z direction and with a magnitude that is proportional to the penetration measured along z from the undulating boundary. Algorithm **F** produces a field resulting from a normal component resulting from penetrating a smooth wall and a lateral component resulting from a friction force modulated as a function of the lateral displacement.

lateral force component which is then modulated according to the height function $h_{\mathbf{F}}(p^x)$. The force field is

$$f_{\mathbf{F}}(p) = \begin{cases} -\kappa_0\left[\mu[1 - h_{\mathbf{F}}(p^x)]\dfrac{d^x}{d^x_{\max}}p^z, \, p^z\right]^\top & \text{if } p^z < 0, \\ [0, 0]^\top & \text{otherwise,} \end{cases}$$
$$h_{\mathbf{F}}(p^x) = \sin(2\pi p^x / L). \tag{3}$$

where the penetration, $d = p^z$, is computed from smooth boundary, d^x is the pre-sliding tangential displacement, $d^x_{\max}$ its the maximum value (1 mm) , μ is Amontons' coefficient of friction. An additional parameter, $d^z_{\max}$ (0.5 mm), is introduced to limit the value of p^z in the computation of the friction force.

The most notable difference lies in the energetic properties of respective force fields given by the algorithms. Neither field is conservative; algorithm **A** produces a generative force field, while **F** gives a field that is generally dissipative.

B. Characteristic number

The algorithms under investigation were thoroughly analyzed in [21]. From this reference, two notions are directly relevant to the present study. The first notion is that if an algorithm generates a non-conservative field, then the resulting synthesis can be tainted by artifacts even if a haptic force feedback system is locally passive. The second is the concept of a characteristic number for a synthesis algorithm. When a virtual wall of stiffness κ_{w} is textured using a given algorithm, the system passivity margin changes according to the norm of the Jacobian matrix of the force field generated by the virtual environment, $\kappa_{\mathrm{t}} = \|J_{\mathrm{ve}}\|_2$. The characteristic number of an algorithm is the ratio $q = \kappa_{\mathrm{t}}/\kappa_{\mathrm{w}}$, which represents the change in passivity margin due to "painting" a texture on a smooth surface. In most cases, the characteristic number is independent from the stiffness of the underlying virtual wall.

C. Experiment Design

The parameters of algorithm **A** were kept fixed when searching for a perceptually equivalent value of parameter μ in algorithm **F**, and only textures having the same spatial frequencies were compared. This procedure avoided the issues related to the possible lack of monotonicity of the relationship spatial-period/roughness. The first step was to define a set

of reference textures to be rendered with algorithm **A**. The characteristic number of this algorithm is of the form $q_{\mathbf{A}} = \sqrt{1 + [2\pi A/L]^2}$ (see appendix for detail). As a consequence, not all the combinations of A and L are admissible. The ratio A/L must then be bounded; for this experiment, $A/L \leq 1$. Selecting five different values for the spatial period, $L = 0.12$, 0.25, 0.50, 1.00, and 2.00 mm gives five values for parameter A as listed in Table I, allowing fifteen valid combinations. This constraint can intuitively be understood by considering the shape of an undulating profile of fixed height. The smaller is the spatial period, the greater becomes the slope, and the greater becomes the 'control gain'.

The stiffness of the virtual wall supporting the texture is $\kappa_{\mathrm{w}} = 1\,000$ N·m^{-1}, hence larger values of $q_{\mathbf{A}}$ indicate cases where non passive behaviors are possible. In particular, for $q_{\mathbf{A}} > 2$, the "effective stiffness" of the texture, $\kappa_t > 2\,000$ N·m^{-1}, would exceed the nominal stiffness of the passive virtual walls achievable with the Pantograph. The characteristic number, however, like other passivity-based measures, is conservative and the parameter value combinations on the diagonal happened to be acceptable in practice. These system constraints caused the experimental design to be unavoidably unbalanced.

TABLE I
CHARACTERISTIC NUMBER $q_{\mathbf{A}}$ FOR THE REFERENCE TEXTURES.

		L (mm)				
		0.12	0.25	0.50	1.00	2.00
	0.12	6.4	3.3	1.9	1.3	1.0
	0.25		6.4	3.3	1.9	1.3
A (mm)	0.50			6.4	3.3	1.9
	1.00				6.4	3.3
	2.00					6.4

The second step was to define the initial intervals for μ to be used with the MOBS thresholding method. The intervals were chosen so that the friction algorithm would be marginally non passive when rendering the highest possible value of μ in the interval, see Table II.

TABLE II
STARTING INTERVALS FOR MOBS ESTIMATION OF μ.

	L (mm)				
	0.12	0.25	0.50	1.00	2.00
top	0.2	0.4	0.6	0.8	1.0
bottom	0	0	0	0	0

For accurate calibration, seven reversals were required and the terminal interval was to be 1% of the size of the inital interval, otherwise the trial was extended for two additional reversals. As already mentioned, a newly introduced termination criterion was needed because, due to passivity constraints, the initial interval could not be made arbitrarily large and therefore was not guaranteed to contain the subjectively equivalent threshold value for μ. To deal with this problem, the trial was ended if the three upper boundaries and the three lower boundaries in the interval stacks were all equal; the threshold was then set to the value of the boundary.

Because intervals cannot be made arbitrarily large, some resulting thresholds values were clipped and did not reflect

a true threshold. Such occurrence was accounted for in the analysis of the results, where order-based tests were preferred. In particular, the median was used to indicate the central tendency of the samples.

D. Subjects and Experimental Procedure

A total of 10 paid subjects participated in the experiment (3 male and 7 female). Although the Pantograph device operates quietly, a faint noise sometimes emanates from the actuators, which may taint the results. Subjects were asked to wear sound isolation headphones (DirectSound™ EX-29) through which white noise was played at a self-adjusted volume. The apparatus was concealed behind a curtain.

Subjects interacted with the Pantograph by putting the index finger of their dominant hand on the Pantograph's plate. An experiment consisted of a sequence of MOBS thresholding trials, three for each of the 15 reference textures. Subjects were presented with two textures synthesized on two parallel facing virtual walls, 30 mm apart. See Figure 3. The reference texture given by algorithm **A** was compared with a test texture given by algorithm **F**.
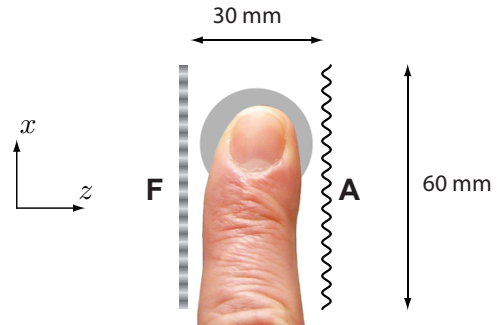

Fig. 3. Experiment layout. The two textured virtual walls were 30 mm apart, randomly disposed on the left or on the right throughout the trials.

Subjects were asked to report which of the two textures was rougher, and their answers were entered via keyboard strokes. After each answer, the state of the MOBS procedure was updated, a new μ was computed and presented to the subject. This sequence was repeated until convergence, which was typically achieved in less than two minutes.

Subjects were given a single training trial, to familiarize them with the haptic device and with the movements required for exploration. No feedback was given. Subjects were instructed to proceed as fast as possible but no restriction was imposed on their exploratory procedure. A stroking mouvement was advised during the training trial. They could explore the surfaces in any order and as many times they felt it necessary to reach a judgment.

Reference textures were presented to each subject in a randomized order. The order was the same for all subjects to facilitate analysis and to avoid bias in the subject's responses. Due to the adaptive nature of the experiment (and hence a variable number of presentations) and because of the lack of exploratory constraints (different exploration time for each of the trials), we did not expect a significant contribution

of learning effects in the resulting thresholds. The reference texture was presented either left or right with a 50% probability. Each subject estimated 3 thresholds for μ with each of the 15 reference textures. At the end of the experiment, 450 thresholds were recorded. To avoid biases, subjects *were not instructed* as to what was meant by roughness. To avoid confusion, "roughness" was simply defined as being "the opposite of smoothness" [31].

IV. RESULTS

A. Raw Data

Subjects tended to produce four patterns of convergence, described as follows.

- Perfect convergence was characterized by an exponentially contracting interval resembling an overdamped second-order response, until the required number of reversals was achieved, see Fig. 4(a).
- 'One-bump' convergence occurred when the threshold drifted sufficiently to force a reset to an initial boundary value. MOBS can still converge after a reset if there is a sufficient number of reversals left in the trial, Fig. 4(b). Some trials exibit multiple bumps before convergence.
- A 'trumpet artifact' occurred when the threshold drifted toward the end of a trial. In response, the interval had to increase dramatically near the end of a trial, see Fig. 4(c). MOBS cannot recover from such ocurrence. These artifacts were observed 1% of the trials and were corrected for by discarding the last 2-3 answers.
- In approximately 5% of all trials, subjects were not able to consistently compare textures, see Fig. 4(d). These cases were not discarded and were considered as noise in the calibration procedure.

Overall, the method was found to converge to a clear threshold, directly or with "bumps", in 94% of all trials. After correcting for the "trumpet" artifacts, the convergence rate was about 95%.

Repeating the thresholding process three times made failure to converge very unlikely. In only one case out of 150 trials two of the three thresholds were not perfectly convergent, but since the two were consistent with each other no action was taken. The few of cases in which MOBS did not converge had no significant influence on the results. The most notable artifact that can be attributed to MOBS result from an exaggerated modification of the intervals in response to drifting thresholds.

B. Analysis Of The Overall Results

Please refer to Fig. 5 for a plot of the distribution of the 450 estimated threshold values sorted by amplitude of the reference texture.

The estimated threshold of μ for the point of subjective equivalence was strongly correlated with the amplitude A in algorithm **F**. The Spearman correlation test gave $\rho = 0.7889$ ($n = 450$, $p < 10^{-6}$) and the Pearson correlation test gave $r = 0.6980$ ($n = 450$, $p < 10^{-6}$), showing a strong and significant monotonic linear correlation between these quantities. Also, the median values of μ (over all subjects
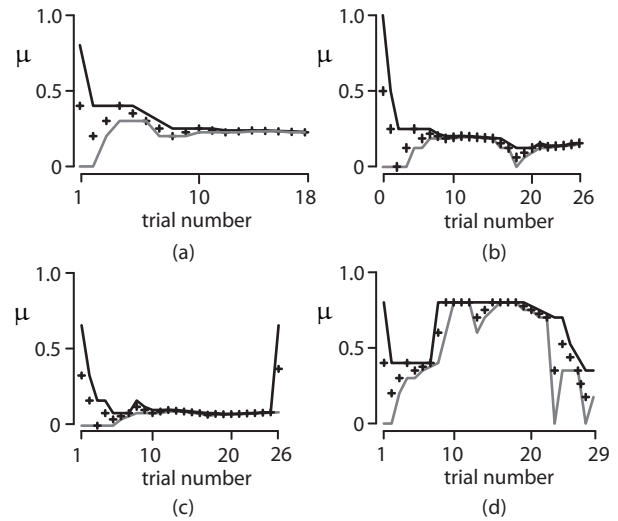

Fig. 4. Types of convergence patterns. The solid line is the upper boundary of the interval. The grey line is the lower boundary. Crosses represent the tested values. (a) Example of exponential convergence. (b) Example of recovery from a drifting threshold. (c) Example of trumpet artifact which occurred in 1% of cases. (d) Rare example of lack of convergence.
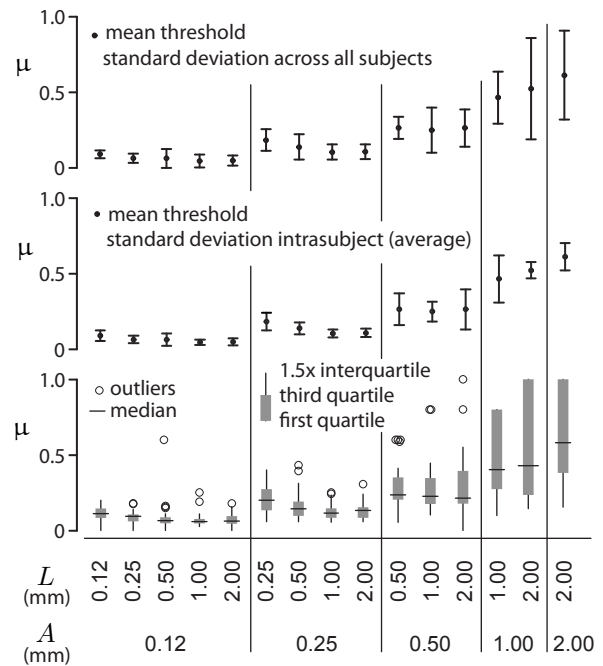

Fig. 5. Results of calibration procedure where 450 thresholds were estimated. The bottom panels shows the statistical distribution of the thresholds. The other panels report the variability within and among subjects.

TABLE III
MEDIANS OF THE ESTIMATED PERCEPTUALLY EQUIVALENT VALUES OF μ.

		L (mm)				
		0.12	0.25	0.50	1.00	2.00
	0.12	0.11	0.09	0.06	0.06	0.06
	0.25		0.20	0.14	0.12	0.13
A (mm)	0.50			0.23	0.23	0.22
	1.00				0.40	0.43
	2.00					0.58

and thresholds) correlated linearly with the amplitude A, $r = 0.9661$ ($n = 15$, $p < 10^{-6}$). Moreover, the transformation A $[\mathbf{A}]$ $\leftrightarrow$ roughness $\leftrightarrow$ μ $[\mathbf{F}]$ was monotonic and largely linear. This strong parameter correlation is an important result and was obtained without explicitly relating the parameters to roughness.

Due to the limited number of data points per reference texture and because of the limited size of the initial intervals, the significance of the data was assessed through non parametric statistics, with Friedman tests in particular. The triangular size of the data forced us to use eight different tests, performed on the estimated thresholds grouped first by spatial frequency and then by amplitude. Subsequently, 20 repeated Friedman tests were perfomed to assess pairwise significant differences; to account for the repeated testing, the significance level was lowered to $\alpha = 2.5 \times 10^{-3}$. Subjects were treated as random row factors with three repetitions, by using a two-way Friedman test, for the significance of the column effect (the parameter A or L) in presence of row effect (the subject).

The first observation regards the distribution of the raw data; fig. 5 shows a large variation among subjects for textures $(A, L) = (1, 2)$ and $(2, 2)$ mm. In addition, the individual differences of the median were statistically significant. When the data were grouped by amplitude, the spatial frequency showed significant effect for $A = 0.12$ mm, ($p < 2 \times 10^{-5}$) and $A = 0.25$ mm, ($p < 2 \times 10^{-4}$). Repeated pairwise Friedman tests suggest that textures with $L \leq 0.25$ mm lead to significantly different estimations of μ than textures with $L > 0.25$ mm. Interestingly, this partition is similar to the distinction made between micro and macro textures [12]. The design of the current experiment, however, and the intrasubject variability for the textures on the diagonal of Table III do not allow to speculate further about this observation.

When the data are grouped by spatial period, we should expect significant differences. In fact, the distribution of μ showed significant differences due to amplitude ($p < 10^{-6}$); the median values of μ for each reference texture are reported in Table III. Further investigations with repeated pairwise Friedman tests showed significant pairwise differences between textures with the same spatial period ($p < 2.5 \times 10^{-3}$). The only exception is the pair $A = 1, 2$ mm, $L = 2$ mm for which significance was not found ($p > 6 \times 10^{-3}$). These results confirm that the calibration procedures generally assigns a significantly different μ estimate for each different A.

Using the results of Table III the characteristic numbers of algorithm $\mathbf{F}$ can be computed and compared to that given by algorithm $\mathbf{A}$, for an equivalent level of roughness sensation. The results are collected in Table IV. By comparing with the values in Table I, algorithm $\mathbf{F}$ can be said to be on average 30% more passive than algorithm $\mathbf{A}$ for an equivalent sensation of roughness. The subjects in this sample were found to be a significant random factor for the estimates of μ, thus the passivity margins should be assesed on a per subject basis, which further motivates the need for a rapid calibration procedure. The passivity margins, computed from the median of the estimates, are nonetheless reported in Table IV.

TABLE IV
CHARACTERISTIC NUMBER OF ALGORITHM **F** FOR EQUIVALENT
PERCEPTUAL ROUGHNESS.

		L (mm)				
		0.12	0.25	0.50	1.00	2.00
A (mm)	0.12	6.10	2.71	1.12	1.01	1.01
	0.25		5.87	2.42	1.25	1.12
	0.50			3.92	2.38	1.63
	1.00				4.23	3.20
	2.00					4.32

V. DISCUSSION AND CONCLUSION

In this paper we have described a method that can be used for fast perceptual calibration between haptic synthesis algorithms. Given careful characterization of the mathematical and control properties of algorithms as well as of the hardware platform used to transduce the signal, it was possible to show that a pair of algorithms operating on different principles could be calibrated to produce an equivalence of roughness for a large range of parameter settings. Efficiency resulted from the use of a fast matching method operating from the principle of dichotomy search, which was adapted to the needs of the determination of the point of subjective equivalence.

Our main results are Tables III and IV. It can be seen from (2) and (3) that each algorithm depends on two parameters where the spatial period, L, is common to both. The amplitude, A, is particular to the algorithm $\mathbf{A}$, while $\mathbf{F}$ depends on the coefficient of friction, μ. It is therefore apparent that the two algorithms operate on different principles — A and μ have different dimensions — yet, they can be perceptually calibrated against each other. Once this result was obtained, we could relate the two algorithms from the view point of the passivity margin that they provide. Specifically, an algorithm based on the principle of lateral friction force modulation, $\mathbf{F}$, was found to be more passive than an algorithm based on perturbing the virtual interaction point in a direction orthogonal to the virtual surface, $\mathbf{A}$. Conversely, algorithm $\mathbf{F}$ can provide a greater sensation of roughness than $\mathbf{A}$ for the same passivity margin.

Although roughness is a perceptual dimension that is frequently employed to discriminate textures, the results of the present study showed an effect of the subject on the perceived equivalent roughness. While the nature of this difference among subjects could not be investigated, the lack of consistency among subjects suggests that algorithm calibration in virtual reality applications need to be calibrated on a per subject basis. The friction coefficient, μ, that drives the generation of non-geometric cues could nevertheless be made to play a role equivalent to that of texture geometric profile amplitude, A. This result was obtained without an explicit estimation of roughness. In respect to the lack of a precise definition for the notion of surface roughness, some subjects interpreted reference textures with large spatial period $((A, L) = ((1, 2), (2, 2))$ mm) as being very rough, while others identified them as very smooth wavy surfaces. Given the precise characterization of the synthesized stimuli, it is anticipated that the data in Table III would also be useful if these algorithms are employed with hardware platforms providing the same level of performance as the Pantograph.

APPENDIX A
CHARACTERISTIC NUMBERS

A. Algorithm **A**

The Jacobian matrix of the force field is

$$\mathbf{J}_{f\mathbf{A}}(p) = -\kappa_0 \begin{bmatrix} 0 & 0 \\ -h'_{\mathbf{A}}(p^x) & 1 \end{bmatrix}. \tag{4}$$

Its norm is

$$\|\mathbf{J}_{f\mathbf{A}}\|_2 = \kappa_{\mathbf{A}} = \kappa_0\sqrt{1 + [h'_{\mathbf{A}}(p^x)]^2} \tag{5}$$

which gives

$$q_{\mathbf{A}} = \max \kappa_{\mathbf{A}}/\kappa_0 = \sqrt{1 + [2\pi A/L]^2} \tag{6}$$

when $h_{\mathbf{A}}(p^x) = A\sin(2\pi p^x/L)$.

B. Algorithm **F**

The Jacobian matrix of the force field is

$$\mathbf{J}_{f\mathbf{F}}(p) = -\kappa_0 \begin{bmatrix} \mu\dfrac{p^z}{d^x_{\max}}\left(\dfrac{\mathrm{d}d^x}{\mathrm{d}p^x} - h_{\mathbf{F}}(p^x)\dfrac{\mathrm{d}d^x}{\mathrm{d}p^x} - h'(p^x)\right) & \mu[1 - h_{\mathbf{F}}(p^x)]\dfrac{d^x}{d^x_{\max}} \\ 0 & 1 \end{bmatrix}. \tag{7}$$

In the worst case and according to [30], (7) becomes:

$$\mathbf{J}_{f\mathbf{F}}(p) = -\kappa_0 \begin{bmatrix} 2\mu p^z(1/d^x_{\max} + \pi/L) & 2\mu \\ 0 & 1 \end{bmatrix} \tag{8}$$

For $h_{\mathbf{F}}(p^x) = \sin(2\pi p^x/L)$, $q_{\mathbf{F}} = \max\|\mathbf{J}_{f\mathbf{A}}\|_2/\kappa_0$ can be numerically evaluated once the value of p^z is clampled to a maximum $d^z_{\max}$.

ACKNOWLEDGMENTS

REFERENCES

[1] G. D. Legge, D. H. Parish, A. Luebker, and L. H. Wurm, "Psychophysics of reading. xi. comparing color contrast and luminance contrast," *Journal of the Optical Society of America, A*, vol. 7, no. 10, pp. 2002–2010, 1990.

[2] M. Stokes, M. D. Fairchild, and R. S. Berns, "Precision requirements for digital color reproduction," *ACM Transactions on Graphics*, vol. 11, no. 4, pp. 406–422, 1992.

[3] P. Seuntiens, L. Meesters, and W. Ijsselsteijn, "Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric jpeg coding and camera separation," *ACM Transactions on Applied Perception*, vol. 3, no. 2, pp. 95–109, 2006.

[4] T. Tolonen and H. Järveläinen, "Perceptual study of decay parameters in plucked string synthesis," in *Proceedings of the 109th Convention of Audio Engineering Society*, 2000, p. preprint no. 5205.

[5] P. Leškovský, T. Cooke, M. O. Ernst, and M. Harders, "Using multidimensional scaling to quantify the fidelity of haptic rendering of deformable objects," in *Proceedings of Eurohaptics*, 2006, pp. 289–295.

[6] V. Hayward and O. R. Astley, "Performance measures for haptic interfaces," in *Robotics Research: The 7th International Symposium*, G. Giralt and G. Hirzinger, Eds. Heidelberg: Springer Verlag, 1996, pp. 195–207.

[7] M. Lin and M. Otaduy, Eds., *Haptic Rendering: Foundations, Algorithms and Applications.* A. K. Peters, Ltd, 2008.

[8] P. P. Ho, B. D. Adelstein, and H. Kazerooni, "Judging 2D versus 3D square-wave virtual grating," in *Proceedings of the 12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems,*, 2004, pp. 176–183.

[9] J. M. Weisenberger, M. J. Kreier, and M. A. Rinker, "Judging the orientation of sinusoidal and square-wave virtual gratings presented via 2-DOF and 3-DOF haptic interfaces," *Haptics-e*, vol. 1, no. 4, 2000.

[10] S. Choi and H. Z. Tan, "Perceived instability of virtual haptic texture. III. effect of update rate," *Presence: Teleoperators and Virtual Environments*, vol. 16, no. 3, pp. 263–278, 2007.

[11] W. M. Bergmann-Tiest and A. M. L. Kappers, "Analysis of haptic perception of materials by multidimensional scaling and physical measurements of roughness and compressibility," *Acta Psychologica*, vol. 121, no. 1, pp. 1–20, 2006.

[12] M. Hollins and S. J. Bensmaïa, "The coding of roughness," *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie experimentale*, vol. 61, no. 3, pp. 184–195, 2007.

[13] A. M. Smith, C. E. Chapman, M. Deslandes, J. S. Langlais, and M. P. Thibodeau, "Role of friction and tangential force variation in the subjective scaling of tactile roughness," *Experimental Brain Research*, vol. 144, no. 2, pp. 211–223, 2002.

[14] M. A. Lawrence, R. Kitada, R. L. Klatzky, and S. J. Lederman, "Haptic roughness perception of linear gratings via bare finger or rigid probe." *Perception*, vol. 36, no. 4, pp. 547–557, 2007.

[15] R. L. Klatzky and S. J. Lederman, "Tactile roughness perception with a rigid link interposed between skin and surface," *Perception and Psychophysics*, vol. 61, no. 4, pp. 591–607, 1999.

[16] S. J. Lederman, R. L. Klatzky, C. L. Hamilton, and G. I. Ramsay, "Perceiving roughness via a rigid probe: Psychophysical effects of exploration speed and mode of touch," *Haptics-E: Electronic Journal of Haptics Research*, vol. 1, 1999.

[17] S. J. Lederman, R. L. Klatzky, C. Hamilton, and M. Grindley, "Perceiving surface roughness through a probe: Effects of applied force and probe diameter," in *Proceedings ASME IMECE Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems*, vol. 69-2, 2000, pp. 1065–1071.

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

[19] M. Hollins, S. Bensmaïa, K. Karlof, and F. Young., "Individual differences in perceptual space for tactile textures: evidence from multidimensional scaling," *Perception & Psychophysics*, vol. 62, no. 8, pp. 1534–1544, Nov 2000.

[20] G. Campion and V. Hayward, "Fundamental limits in the rendering of virtual haptic textures," in *Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 2005, pp. 263–270.

[21] ——, "On the synthesis of haptic textures," *IEEE Transactions on Robotics*, vol. 24, no. 3, pp. 527–536, 2008.

[22] G. Campion, A. H. Gosline, and V. Hayward, "Does judgement of haptic virtual texture roughness scale monotonically with lateral force modulation?" in *Proceedings of Eurohaptics*, vol. LNCS 5024. New York: Springer-Verlag, 2008, pp. 718–723.

[23] G. Campion, Q. Wang, and V. Hayward, "The Pantograph Mk-II: A haptic instrument," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS'05*, 2005, pp. 723–728.

[24] A. M. Smith, G. Gosselin, and B. Houde, "Deployment of fingertip forces in tactile exploration," *Experimental Brain Research*, vol. 147, pp. 209–218, 2002.

[25] A. H. C. Gosline and V. Hayward, "Eddy current brakes for haptic interfaces: Design, identification, and control," *IEEE/ASME Transactions on Mechatronics*, vol. 13, no. 6, pp. 699–677, 2008.

[26] J. E. Colgate and G. Schenkel, "Passivity of a class of sampled-data systems: Application to haptic interfaces," in *Proceedings of the American Control Conference*, 1994, pp. 3236–3240.

[27] C. Kaernbach, "Slope bias of psychometric functions derived from adaptive data," *Perception and Psychophysics*, vol. 69, no. 8, pp. 1389–1398, 2001.

[28] R. A. Tyrrell and D. A. Owens, "A rapid technique to assess the resting states of the eyes and other threshold phenomena: the modified binary search (MOBS)," *Behavior Research Methods, Instruments, and Computers*, vol. 20, no. 2, pp. 137–41, 1988.

[29] A. J. Anderson and C. A. Johnson, "Comparison of the ASA, MOBS, and ZEST threshold methods," *Vision Research*, vol. 46, no. 15, pp. 2403–2411, 2006.

[30] V. Hayward and B. Armstrong, "A new computational model of friction applied to haptic rendering," in *Experimental Robotics VI*, ser. Lecture Notes in Control and Information Sciences, P. Corke and J. Trevelyan, Eds., vol. 250, 2000, pp. 403–412.

[31] J. C. Stevens and J. R. Harris, "The scaling of subjective roughness and smoothness," *Journal of Experimental Psychology*, vol. 64, pp. 498–494, 1962.

Gianni Campion received his M.Sc. in Computer Science from University of Verona (Italy) in 2000. He joined the Haptics Laboratory at McGill University in 2002, where he is pursuing a Ph.D. in Electrical Engineering. His research interests include force feedback displays, surgical robotics, and haptic textures.

Vincent Hayward (M'84-SM'04-FIEEE'08), Ing. Ecole Centrale de Nantes 1978; Ph.D. Computer Science 1981, University of Paris; Visiting Assistant Professor, Purdue University (1982); Chargé de Recherches at CNRS, France (1983-86), Professeur invité, Université Pierre et Marie Curie, UPMC (2006); is Professor of Electrical and Computer Engineering at McGill University and has recently taken up a chaired position at UPMC. Hayward is interested in haptic device design and applications, perception, and robotics. He is leading the Haptics Laboratory at McGill University since 1994 and was the Director of the Center for Intelligent Machines (2001-2004). He is a co-Founder of the Experimental Robotics Symposia, Program Vice-Chair 1998 IEEE Conference on Robotics and Automation, Program Vice-Chair ISR2000, past Associate Editor of the IEEE Transactions on Robotics and Automation, now member of the Governing board of Haptics-e, of the Editorial board of the ACM Transactions on Applied Perception, and of the IEEE Transactions on Haptics. Hayward received best paper and research awards including the NASA Space Act Tech Brief Award (1991) and the E. (Ben) & Mary Hochhausen Award for Research in Adaptive Technology For Blind and Visually Impaired Persons (2002).