Stiffness in Virtual Contact Events: A Non-Parametric Bayesian Approach

Jonathan Browder, Séréna Bochereau, Femke van Beek, Raymond King

***Abstract*— In this study we investigated the use of simple vibrotactile signals to simulate contact with a virtual object. In particular we explored the relation between properties of the signal and the perceived hardness of the object. The space of stimuli is large, and we have no plausible a priori model for the relationship of parameters to percept. Thus we made use of non-parametric Bayesian methods, in particular utilizing Gaussian process priors. We show that this method both gives insight into the phenomenon of interest and well-predicts a second, separate data set collected via the method of constant stimuli. Thus we argue that it could be a fruitful approach for attacking a variety of perceptual problems.**

I. INTRODUCTION

Within the haptic domain, how hard an object is perceived to be is driven by many different factors (force, position, stiffness, damping, etc.), in a complex manner [1], [2]. Visual cues can also be highly relevant, [3], [4]. A popular method for simulating contact with virtual objects, for instance for use in VR applications, is the use of vibrations at the moment of contact. Several groups have previously investigated this phenomenon, e.g.; Kuchenbecker [5] and Okamura [6] used vibration signals built from mechanically motivated models to augment the feeling of contact through a stylus with force feedback, and showed that these improved the realism of virtual contact.

In this study, we investigate similar vibratory signals, but we used them to simulate passive contact directly on the fingertip. We did no provide force feedback, but did use high-quality virtual reality visuals: a virtual sphere falling onto the upturned finger. While previous studies have derived vibration parameters from measurements of real vibrations resulting from contact with a small set of objects of varying stiffness, we attempted to directly map the relationship of vibration properties to their perceptual consequences. In particular, we studied the perception of hardness, over a large portion of the space of possible parameter values.

We took decaying sinusoids as in [6], determined by frequency f and decay constant d, as our vibrotactile cues. To determine how these parameters relate to the perception of hardness of a virtual object, we presented participants with two stimuli and asked them to report which one was 'harder' using a two-alternative forced-choice (2AFC) task.

In the classic psychophysical paradigm, a 'reference' or 'standard' stimulus would be fixed, and compared against a collection of 'comparison' stimuli. Often the next step is to fit a psychometric function, of the form

$$P(\text{“harder than reference”}) = \Psi(aX+b), \tag{1}$$

where Ψ is some sigmoid curve (generally a cumulative density function), and X is some measure of the 'hardness' of the comparison, or at least a parameter of the comparison stimulus that we expect to be linearly related to hardness (perhaps after a transformation such as taking the logarithm) [7].

However, we have no reason to expect perceived hardness to be even monotonically related to the parameters of our vibrotactile signals, let alone linearly related. Further, this method allows us only to relate a single point in stimulus space (the standard) to a one-dimensional subspace (the comparisons parametrized by X); mapping out any significant portion of our 2-dimensional stimulus space in this manner would require a very large amount of data.

Thus we instead take an approach in which we test many different pairs of stimuli, and model a single function that determines the probability of the judgment between any two stimuli:

$$P(\text{“A harder than B”}) = \Psi(F(A)-F(B)), \tag{2}$$

where F is a latent function we wish to estimate. We we do not assume F to be linear. Instead, we use Gaussian process priors to model F, which allows us to learn the shape of F while enforcing only weak assumptions on that shape.

We believe this method has wide applicability in the field of haptic perception, since there are many areas where determining mappings of multi-dimensional stimulus spaces to haptic percepts would be of great interest, but models for a mapping are lacking (especially when working with artificial stimuli).

Thus we begin with an explanation of our framework in the general case of modeling two-alternative forced choice data in a Bayesian manner and fitting with Gaussian process priors. We then demonstrate the use of the method on an experimental data set investigating virtual contact (described herein). We validate the approach by comparing it to a second experimental data set; using the same task but with data collected via the method of constant stimuli — comparing a set of comparison stimuli to a fixed reference, with each comparison repeated several times. We show that this data is well predicted by the non-parametric model.

Facebook Reality Labs, Redmond Washington
jonathan.browder@oculus.com

II. MODELING

A. Set Up

We will suppose for the following that we have collected data by presenting participants with two stimuli, S and S', and requiring the participant to choose one over the other on some criteria (e.g., which is harder). We note, however, that the general method is useful, with appropriate modifications, for a wide variety of applications.

We will assume that our stimuli can be parametrized as real vectors, $S, S' \in \mathbb{R}^D$ (some of these could be categorical variables encoded as real numbers). We will denote by $S \succ S'$ the event that S is chosen over S', so our data takes the form of a set of n such relations, one for each trial: $\mathscr{D} = \{S_i \succ S_i'\}_{i=1}^n$. We are interested in estimating the probability of choices dependent on the values of the stimuli, $P(S \succ S')$, which we will model by

$$P(S \succ S') = \Phi(F(S) - F(S')), \tag{3}$$

where Φ is the cumulative distribution function of $N(0,1)$ and F is an unknown 'latent function' which must be estimated. To arrive at this model (which could also simply be accepted on purely pragmatic grounds), consider the following:

Imagine that each stimulus S has some perceived 'amount' of the property on which it is being judged, which we call $F(S)$. However, noise in the perceptual system means that what is actually perceived when S is presented is

$$\hat{F}(S) = F(S) + \varepsilon, \tag{4}$$

where $\varepsilon \sim N(0, \sigma^2)$. When presented with two stimuli, a participant makes the judgment by comparing $\hat{F}(S)$ and $\hat{F}(S')$, so that $S \succ S'$ exactly when $\hat{F}(S) > \hat{F}(S')$. Then

$$\begin{aligned} P(S \succ S') &= P(0 < \hat{F}(S) - \hat{F}(S') \\ &= P(0 < F(S) + \varepsilon - F(S') - \varepsilon') \\ &= P(\varepsilon' - \varepsilon < F(S) - F(S')) \\ &= P\left(\frac{\varepsilon' - \varepsilon}{\sqrt{2}\sigma} < \frac{F(S) - F(S')}{\sqrt{2}\sigma}\right) \\ &= \Phi\left(\frac{F(S) - F(S')}{\sqrt{2}\sigma}\right), \end{aligned}$$

since $\frac{\varepsilon' - \varepsilon}{\sqrt{2}\sigma} \sim N(0,1)$. Now, since only the relative values of F are meaningful for the model, we can replace F with the equivalent function $\frac{F}{\sqrt{2}\sigma}$. In other words, we require that the 'units' of F be standard deviations of the noise on a single comparison. Our model then reduces to (3).

While this is a useful and generally plausible model, one should be cognizant of its consequences and limitations. One is 'symmetry': it is not aware of any ordering of the stimuli (say, that one is presented first, or that they are presented at different hands). Another is a kind of transitivity: for any S, S' and S^*, $P(S \succ S')$ and $P(S' \succ S^*)$ completely determine $P(S \succ S^*)$, since $F(S) - F(S^*) = F(S) - F(S') + F(S') - F(S^*)$.

If one suspects these properties fail in a significant way the model should be modified. For example, if you think there may be a bias due to the order of stimulus presentation, you may add a term for this in the model, considering the pairs now as ordered.

B. A Bayesian Approach

The question now is how we should estimate F. In principal F could be any function $\mathbb{R}^D \to \mathbb{R}$, but without any further assumptions we will grossly overfit. One approach is to require F to have a certain parametric form, so that fitting F requires only estimating a finite set of parameters; for example we might assume it is linear in S, so that the parameters are the linear coefficients (if we then require that there is a 'standard' stimulus S that appears in each pair in the data set, and $D = 1$, we reduce to the familiar psychometric curve common in the literature). But there are cases when we do not have good, theory driven models for F. To attack these, we propose a non-parametric Bayesian approach.

Recall that by Bayes's theorem,

$$P(F|\mathscr{D}) \propto P(\mathscr{D}|F)P(F). \tag{5}$$

Assuming our trials are independent, we obtain from (3) our likelihood

$$P(\mathscr{D}|F) = \prod_{i=1}^{n} \Phi(F(S_i) - F(S_i')). \tag{6}$$

It is through our prior, $P(F)$, that we will impose structure on our functions. To do so, we model $P(F)$ as a *Gaussian Process* (GP). (For further details and a much broader exploration of Gaussian processes, see, e.g., [8].)

Consider the set of values of F as an infinite collection of random variables, $\{F(S)\}_{S \in \mathbb{R}^D}$. Such a collection is defined to be a *Gaussian process* if for any finite subset $S_1, S_2, \ldots, S_k \in \mathbb{R}^D$, the joint distribution of $F(S_1), F(S_2), \ldots, F(S_k)$ is Gaussian:

$$\begin{bmatrix} F(S_1) \\ F(S_2) \\ \vdots \\ F(S_k) \end{bmatrix} \sim N\left([\mu(S_i)]_i, [K(S_i, S_j)]_{i,j}\right),$$

where the vector $[\mu(S_i)]$ is the mean vector and the matrix $[K(S_i, S_j)]_{i,j}]$ is a covariance matrix; the entries of these depend on the points S_i, and thus may be considered as functions $\mu : \mathbb{R}^D \to \mathbb{R}$ and $K : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}$, respectively. We denote this Gaussian process $GP(\mu, K)$.

For the discussion that follows, we will take $\mu \equiv 0$ (for reasons we will see, this is a reasonable choice in most of our applications). The function K is called either the covariance function or kernel, and must be chosen so that for any selection of points S_i, the resulting matrix is positive semi-definite. This function tells us the (prior) covariance between the values of the function F at various points, and thus encodes our assumptions on the shape of F.

One popular, and useful, choice of covariance function is the square-exponential covariance function,

$$K(S, S') = \alpha^2 e^{-\sum_{j=1}^{D} \frac{(S^j - S'^j)^2}{2\rho_j^2}},$$

where S^j and S'^j denote the j^{th} components of S and S', respectively. Note that under this covariance function, the covariance between $F(S)$ and $F(S')$ is large when S and S' are close, and decreases exponentially as the distance between them increases. Thus, F is likely to take similar values at points that are close together. In fact, draws from a GP with this covariance function will be smooth (infinitely differentiable) with probability 1. This makes it a good choice for many applications (though there are a variety of covariance functions that can be used, some with weaker smoothness assumptions that may be more appropriate to an application).

The constant α determines the 'magnitude' of draws from the GP (note that $K(S,S) = \alpha^2$, so α^2 is the variance of single values about the mean).

The constant ρ_j determines the rate with which the covariance declines with increase in distance in the j^{th} component of S, and thus how quickly the function varies in the j^{th} direction. The larger ρ_j, the less functions drawn from the Gaussian process vary with the j^{th} component of S, and thus the less relevant this component is to determining the value of F. Thus the process of learning ρ_j from the data is sometimes called 'automatic relevance determination.'

The function F is determined by infinitely many values (one for each possible S), but note that the likelihood only depends on the value of F at the sampled data points. Thus, in practice, when we estimated the posterior for F, what we do is select a finite set of points $S_1, S_2, \ldots, S_m$ that includes all of our sampled points (and any other points we wish to investigate), and take the posterior of F evaluated at these points:

$$P([F(S_i)]_{i=1}^m|\mathscr{D}) \propto P(\mathscr{D}|[F(S_i)]_{i=1}^m)P([F(S_i)]_{i=1}^m) \quad (7)$$

where the likelihood is again given by (6) and, by the definition of a Gaussian process, $P([F(S_i)]_{i=1}^m)$ is a multivariate Gaussian,

$$P([F(S_i)]_{i=1}^m) \sim N(0, [K(S_i,S_j)]_{i,j}). \quad (8)$$

To estimate our length-scale and magnitude parameters from the data, we include priors for them, so our full Bayesian model becomes

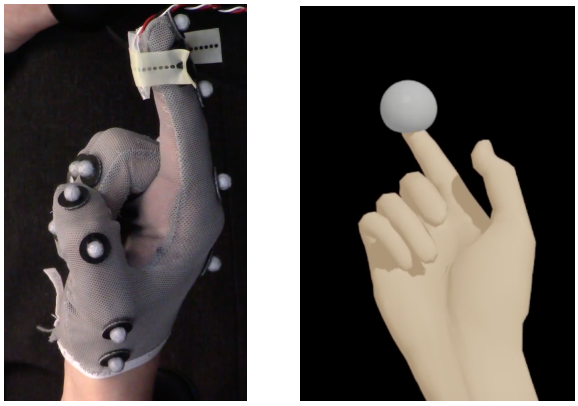
$$P(F|\mathscr{D}) \propto P(\mathscr{D}|F)P(F|\alpha,\rho)P(\alpha,\rho). \quad (9)$$

III. Vibration for Contact

A. Methods

We had 17 healthy participants, aged 18-61, 8 male and 9 female. All but one were right-handed. All participants gave written informed consent prior to taking part in the study and were naïve to the purpose of the experiment. None of them had any history of neurological disorders. All experimental protocols were approved by WIRB.

Participants wore a head mounted display (the Oculus Rift) and sat with their right arm set in an arm rest, their palm facing up and their right index outstretched. They wore a powermesh glove with fidicuals, allowing their hand to be tracked by OptiTrack motion tracking cameras and to be rendered in virtual reality co-localized with their real hand. A commericially available TDK Piezo Haptic Actuator — PowerHap 7G Type (12mm x 12mm x 1mm), encased in a silicone layer — was attached to the tip of their right index finger, with minimal pressure, using an adjustable latex strap, Figure 1a. A pressure sensor (Tekscan FlexiForce ESS301 Sensor) was used to ensure the pressure was adequate. Pink noise was played in the participant's ears through the Rift, to ensure they could not hear the motor vibrations. A virtual gray ball was shown dropping onto the participant's finger (Figure 1b), and at the moment of contact a short signal was played. On each trial, two ball drops were presented in sequence, and the participant was asked "Which sphere was Harder?", and responded using an Oculus touch controller held in their other hand. Across all trials, the ball displayed had the same, slightly soft, NVIDIA FleX physics engine properties, but the vibrations generated at contact were different.



(a) Subjects wore an optically tracked fiducial glove with vibrotactor strapped over the fingertip.

(b) Virtual Ball Dropping on Finger in VR. The ball deforms slightly as it makes contact.

Fig. 1

Our signals were simple decaying sinusoids, of the form

$$Ae^{-dt}\sin(2\pi ft),$$

where t is time and A, d and f, the amplitude, decay constant and frequency, respectively, are adjustable parameters. To avoid perceived intensity contributing to participant responses, we first ran a short study in which we used two interleaving staircases to determine at which amplitudes the various frequency and decay constant combinations were perceived to have approximately equal intensity. We used those values of A when presenting each (d,f) pair in our experiment (Table 2).

We sampled at 108 distinct stimuli, represented as two-dimensional vectors (d,f). Values of d ranged from 10 to 200 s^{-1} and values of f from 25 to 250 Hz. Due to limitations of the vibrotactor we could not sample a full grid in this range (at certain values the vibrotactor could not reach sufficient amplitudes). The sampled points are plotted over the heatmap in Figure 3a. Each of the 17 participants was

Decay	20	60	100	120	170	200
10 Hz	21	168	967			
25 Hz	1.45	6.43	11.42			
50 Hz	1.36	3.01	4.66			
75 Hz	0.32	2.24	4.16	10.36		
100 Hz	0.81	1.70	2.58	6.81	17.55	
150 Hz	0.24	1.51	2.78	5.31	12.78	
200 Hz	0.37	1.24	2.10	3.90	11.73	16.42
250 Hz	0.24	1.09	1.93	4.10	13.25	18.70

Fig. 2: Amplitudes (V) for stimuli at various frequency and decay constants. Values for decay constants not shown are linearly interpolated from shown values at the corresponding frequency.

presented with a random selection of pairs of vectors, chosen so that each vector (d, f) was displayed to each participant in exactly two trials, once presented first and once presented second.

In the same session we collected an additional data set from the same 17 participants, to test against our Gaussian process model. In these trials, we fixed a reference stimulus at $d = 100\ s^{-1}$, $f = 150$ Hz, roughly in the center of the space. We had participants compare this stimulus to 8 points on the line $f = 150$ Hz and eight points on the line $d = 100\ s^{-1}$. Each trial was presented 14 times per participant, in random order, with the comparison presented first on half the trials and second on the other half.

B. Model and Fitting

We modeled our data as in (9), with likelihood as in (6), using as a prior for F a Gaussian process with mean 0 and square-exponential covariance function

$$K((f,d),(f',d')) = \alpha^2 e^{-\frac{(f-f')^2}{2\rho_f^2} - \frac{(d-d')^2}{2\rho_d^2}}.$$

Recall that hyperparameter α controls the magnitude of draws from the process, while ρ_f and ρ_d determine how quickly functions drawn from the prior vary with f and d, respectively. We used a $N(0,1)$ prior for α and an inverse Gamma prior for both ρ_f and ρ_d. Before fitting we scaled both f and d to have mean 0 and standard deviation 1, thus ensuring our scale priors covered a reasonable range of values.

We coded our model in the Bayesian modeling language Stan [9], [10], and used its built-in MCMC sampler to obtain 4000 draws from our posterior for F, at a 19×11 grid of values including our sampled points. This posterior was fit only using data from our first data set.

Note that F is unidentifiable; replacing F with $F + k$ for some constant k does not affect response probabilities, since $\Phi((F(S)+k) - (F(S')+k) = \Phi(F(S) - F(S'))$ — this why we may safely set the mean of the prior to 0. We are only interested in relative values of F. Thus we subtracted from each draw of F its value at the central point $d = 100$, $f = 150$, so that all values of the latent function are expressed relative to the value at this point.

We compared the results of our fit to the response proportions in our second data set, that is, the proportion of the time each comparison was judged as 'harder' than the reference.

To make this comparison, we calculate the (point-wise) posterior response quantiles from our posterior, observing that, letting M be the number of 'harder' responses at a particular level,

$$P(M \leq K) = \sum_{m=0}^{K} \int P(M = m|F) P(F|\mathscr{D}) df. \qquad (10)$$

Recall $P(F|\mathscr{D})$ is the posterior distribution for F, given our Gaussian process prior and our data. Since we had 238 trials at each level, for level comparison level S we have

$$P(M = m|F) = \binom{238}{m} p^m (1-p)^{238-m}, \qquad (11)$$

where $p = \Phi(F(S) - F(150, 100))$. We approximate the integral in each term of 10 by computing this value for each of our 4000 draws from our posterior for F and taking the mean. We then compute the cumulative sums and plot 80% and 95% intervals (those between the .1 and .9 quantiles and the .025 and .975 quantiles, respectively), along with the median response (all scaled to proportions by dividing by number of trials at a level).

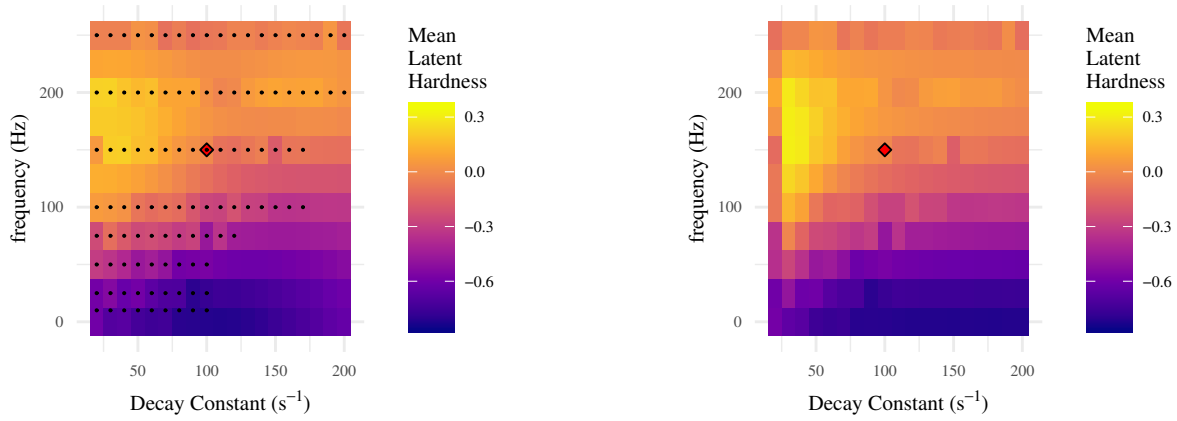
IV. Results and Discussion

We computed the posterior mean of the latent 'hardness' function at each point in our grid; a heat map of the results is shown in Figure 3a.

The latent function is fairly flat, meaning differences in perceived hardness were not large across most of the stimulus space, but shows some interesting trends: stimuli seem to be felt as most stiff when the frequency is around 200 Hz and the decay constant is in the 40-50 s^{-1} range. The softest stimuli occur when the frequency is low.

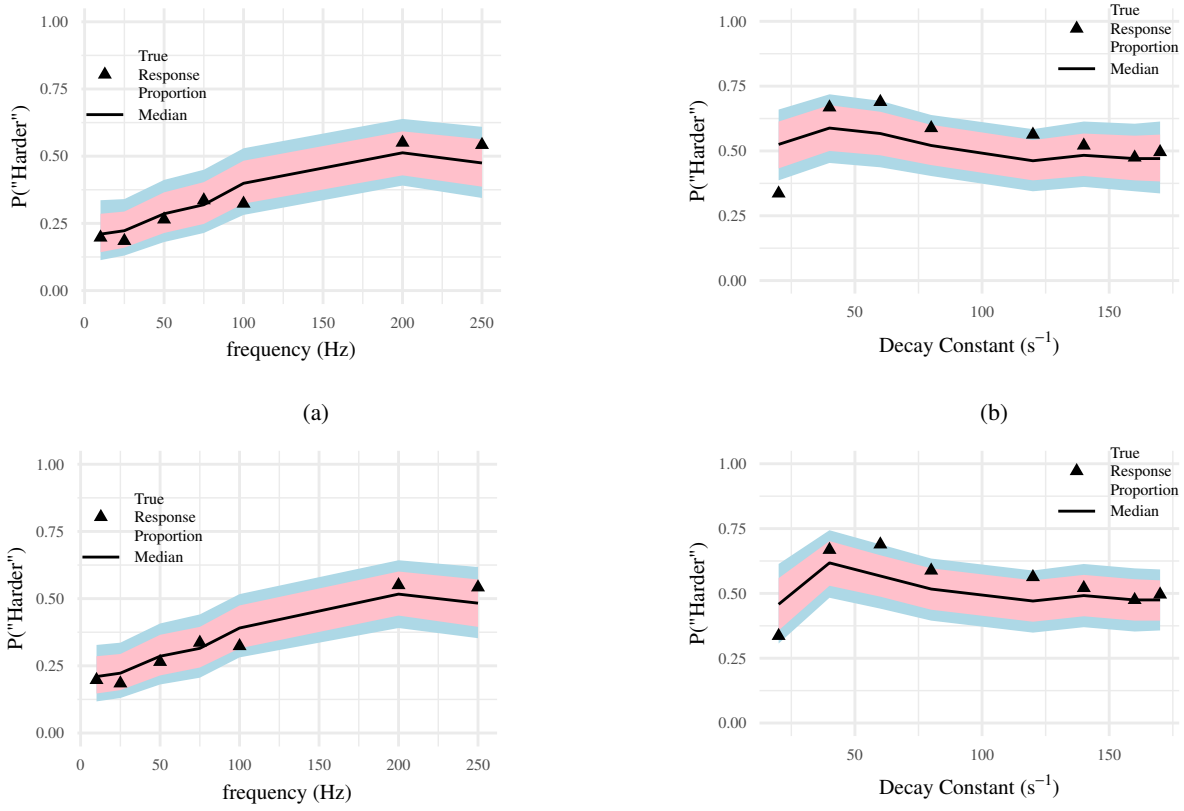
Because we have draws from a full posterior, it is also simple to quantify the uncertainty of our predictions (up to the assumptions of the model). For example, looking solely at the mean of the posterior, we seem to see dips in the latent function value from $(30, 150)$ to $(20, 150)$, and in fact our posterior, even in the original model, has $F(30,150) > F(20,150)$ 89.5% of the time (we calculate this simply be looking at the proportion of the 4000 draws in which this ordering occurs), so we can be somewhat confident that this trend really does exist in our data. On the other hand, we would also see a small dip from $(100, 225)$ to $(100, 250)$, but from our draws we only have around a 74% probability of this ordering, so we are rather less confident (and indeed in our second data set we will not see this trend).

1) Comparison to Second Data Set: We compare our second data set to the posterior response quantiles predicted by our Gaussian process model, Figures 4a and 4b. Note that the plots show point-wise intervals, but the values of F in our posterior are in fact highly correlated, so that for example if we see response proportions in the low end of the interval at



(a) (b)

Fig. 3: Heat map of the mean of the posterior for the latent 'hardness function', relative to the value at $(100\ s^{-1}, 150\ \text{Hz})$ (red diamond). Left: Fit with covariance stationary in d and f, white points show the locations of stimuli present in the data set. Right: Fit with covariance stationary in $\frac{1}{d}$ and f. Values of F determine probabilities of stimuli being deemed harder than one another, so if the value at (d, f) is greater than that at (d', f') by 1, the probability of (d, f) being chosen as harder than (d', f') is $\Phi(1) \approx .84$, if the difference is 1.5 the probability is $\Phi(1.5) \approx .93$, etc.



(a) (b)

(c) (d)

Fig. 4: Median posterior predictive response proportions when comparing against $(100\ s^{-1}, 150\ \text{Hz})$, as predicted by GP model, with central intervals (80% in pink, 95% in light blue), plotted against actual response proportions. Left column: Predictions on the line $f = 150$ Hz. Right Column: Predictions on the line $d = 100\ s^{-1}$. Top row: Predictions from model with prior covariance stationary in f and d. Bottom row: Predictions from model with prior covariance stationary in f and $\frac{1}{d}$.

one point, it is more likely that we will see low proportions also at neighboring points. We see that our Gaussian Process model has done a pretty good job of predicting our second data set. At the same time, Figure 4b reveals something interesting: the sharp dip from $d = 40\ s^{-1}$ to $d = 20\ s^{-1}$. Our fit reflects this dip, which is good, but shows it is as substantially smaller than what we see in the second data set. The Gaussian process has perhaps over-smoothed the effect.

In fact, it is easy to guess why. The square exponential covariance function is what we call 'stationary': the covariance between $F(S)$ and $F(S')$ depends only on $S - S'$. This means the function should vary at the same rate with d and f everywhere in stimulus space. In many contexts this is a safe assumption (even if we do not exactly believe it to be true, small violations should be overcome by the data). But here we might expect it to be violated rather badly: We have expressed our points in (d, f) space, so the covariance between, say, $(20\ s^{-1}, 150$ Hz) and $(40\ s^{-1}, 150$ Hz) is the same as between $(180\ s^{-1}, 150$ Hz) and $(200\ s^{-1}, 150$ Hz). But the duration of a signal (defined as time to fall below a fixed amplitude threshold) is proportional to $\frac{1}{d}$. So a signal with $d = 20\ s^{-1}$ is twice as long as one with $d = 40\ s^{-1}$, while a signal with $d = 180\ s^{-1}$ is only 1.1 times as long as a signal with $d = 200\ s^{-1}$ (and also in absolute terms the difference in duration will be much smaller in the latter case). So we should not be surprised to see F vary more quickly with d when d is small.

With this in mind, we fit a new posterior, replacing d with $\frac{1}{d}$. Results (plotted in the original d for comparison) are shown in figures 4c, 4d and 3b. As expected, our new fit shows a sharper behaviour at low values of d (very long stimuli appear to be experienced as softer), while the modest changes with f are not effected. This illustrates the importance of posterior checks, and careful examination of even the modest assumptions of in our model.

2) Further directions: These results should not be taken as the end of the story, but as a guide towards future investigations. For example, it seems it would be wise to investigate the dip with small d in more detail. The shape of the latent function could also be taken as a starting point for hypothesizing parametrized models.

We also note that the analysis above is only the tip of the iceberg in terms of what can be done with Gaussian processes. For example, different covariance functions can encode different assumptions about the shape of F (perhaps encoding periodicity or underlying monotonic trends). We can also take different likelihoods; for example we get an actually simpler model if the probability of our binary outcome is $\Phi(F(S))$. This could be used when we have a single yes/no outcome dependent on a single stimulus, or two stimuli with a fixed ordering encoded as a single vector, which would allow the above analysis without the symmetry condition.

Finally, we note that in addition to quantifying uncertainty of inferences, possession of a full posterior allows the development of efficient or adaptive sampling strategies that attempt to find the most informative points at which to collect new data.

V. Conclusions

Bayesian techniques with Gaussian process priors can be a useful tool for exploring perceptual data where the underlying relationship between stimuli and responses has unknown form. In our experiment, it reveals a distinct relationship between vibration parameters and perceived hardness of virtual contact, which, while not large in magnitude, is borne out by a posterior check against a second data set. In large portions of the space participants seem largely insensitive to changes in d and f, with values of the latent function varying only slightly. Still, we appear to see peak hardness (from a decaying sinusoid) when the frequency is around 200 Hz and the decay constant is relatively small ($\sim 40\ s^{-1}$).

These methods could be applied to a range of haptic phenomenon; for example we could expand the current study by integrating visual hardness features as further parameters of our stimuli. We hope to pursue these lines of inquiry in future research.

Acknowledgments

Thanks to Max di Luca, Cesare Parise, Elia Gatti, and Majed Samad for helpful advice and suggestions on the design of the haptic experiments, and Andrew Doxon for software support.

References

[1] F. E. van Beek, W. M. Bergmann Tiest, A. M. L. Kappers, and G. Baud-Bovy, "Integrating force and position: testing model predictions," *Experimental Brain Research*, vol. 234, no. 11, pp. 3367–3379, 2016.

[2] F. E. van Beek, D. J. Heck, H. Nijmeijer, W. M. B. Tiest, and A. M. Kappers, "The effect of global and local damping on the perception of hardness," *IEEE transactions on haptics*, no. 3, pp. 409–420, 2016.

[3] M. Korman, K. Teodorescu, A. Cohen, M. Reiner, and D. Gopher, "Effects of order and sensory modality in stiffness perception," *Presence: Teleoperators and Virtual Environments*, vol. 21, no. 3, pp. 295–304, 2012.

[4] M. Kuschel, M. Di Luca, M. Buss, and R. L. Klatzky, "Combination and integration in the perception of visual-haptic compliance information," *IEEE Trans. Haptics*, vol. 3, no. 4, pp. 234–244, 2010.

[5] K. J. Kuchenbecker, J. Fiene, and G. Niemeyer, "Improving contact realism through event-based haptic feedback," *IEEE Trans. Visual. Comput. Graphics*, vol. 12, no. 2, pp. 219–230, 2006.

[6] A. M. Okamura, M. R. Cutkosky, and J. T. Dennerlein, "Reality-based models for vibration feedback in virtual environments," *IEEE/ASME Trans. Mechatron.*, vol. 6, no. 3, pp. 245–252, 2001.

[7] N. Prins *et al.*, *Psychophysics: a practical introduction*. Academic Press, 2016.

[8] C. Rasmussen and C. Williams, *Gaussian Processes for Machine Learning*, ser. Adaptative computation and machine learning series. University Press Group Limited, 2006. [Online]. Available: https://books.google.com/books?id=vWtwQgAACAAJ

[9] Stan Development Team. (2018) Rstan: the R interface to stan, R package version 2.17.3. [Online]. Available: http://mc-stan.org

[10] B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell, "Stan: A probabilistic programming language," *Journal of statistical software*, vol. 76, no. 1, 2017.