A Kalman filter model for predicting performance in natural haptic explorations

Anna Metzger$^1$ and Knut Drewing$^1$

Abstract—Haptic exploration of objects usually consists of repeated exploratory movements and our perception of their properties is the result of the integration of information gained during each of these single movements. The serial nature of information integration in haptic perception necessitates that perceptual estimates from single exploratory movements are retained in memory. Here we propose an optimal model for serial integration of information in haptic perception considering memory limitations by extending a previously proposed model to more natural explorations. We test the model by predicting discrimination performance in free and restricted explorations and the magnitude of performance in free explorations, suggesting that integration of information at least in natural haptic exploration can be well approximated by optimal integration given memory limitations. We further used the predictions of the model to assess whether participants adjust their exploration to maximize performance, by e.g. using an optimal switching strategy. Our results suggest that given the trade of between the improvement of performance and muscular costs of switching, participants invest additional muscular costs only as long there is maximal gain in performance.

I. INTRODUCTION

When we explore objects with the sense of touch, sometimes a single glance is sufficient, for instance to recognize a known object [1] but exploration of some object properties such as softness or texture usually consists of repeated highly stereotypical movements [2]. For instance people repeatedly indent an object or squeeze it between the fingers if they want to perceive its softness or they move the hand back and forth sideways over its surface in order to perceive its roughness. It was shown that such extension of the exploration over space and time increases perceptual reliability [3], [4], [5], [6], [7], consistent with the model of optimal integration of information: Maximum Likelihood Estimation (MLE; [8]). This model predicts that an estimate (i.e., the percept) of information: Maximum Likelihood Estimation (MLE; [8]). This model predicts that an estimate (i.e., the percept) of information: Maximum Likelihood Estimation (MLE; [8]). This model predicts that an estimate (i.e., the percept) of information: Maximum Likelihood Estimation (MLE; [8]). This model predicts that an estimate (i.e., the percept) [9], [10]:

$$\hat{S} = \sum_{i=1}^{n} w_i \hat{s}_i, w_i = \frac{\sigma^{-2}}{\sum_{i=1}^{n} \sigma_i^{-2}}, R_S = \sum_{i=1}^{n} r_i$$  \hspace{1cm} (1)

with the weights constrained to be between 0 and 1 and to sum to 1. While the MLE model correctly predicts the general increase of perceptual reliability with the increasing number of available information it overestimates the amount of the increase achieved with the extension of the exploration in haptic perception [6], [7]. Our recent results are consistent with the idea that the improvement of reliability by increasing the extension of the explorations is limited by memory [11], [12], [13], [14]. In contrast to other perceptual situations such as multisensory or cue integration which are well described by the MLE model, in haptic perception sensory estimates are not simultaneously available but acquired serially. This implies that previously acquired estimates need to be stored in memory especially in the case of a perceptual comparison. Indeed, when the softness or the texture of two objects is compared with the sense of touch latter indentations of the second stimulus or strokes over its surface contribute less to the comparison, consistent with the idea that the previously acquired representation of the first stimulus fades in memory [7], [12], [13], [14].

To account for the serial nature of estimation and memory limitations in haptic perception a modified Kalman filter model was proposed [13]. The Kalman filter is an optimal model for estimating the state of dynamic linear systems over time [15]. Serial integration of information is modeled there as recursive combination of the current estimate with the information obtained from prior estimates considering the dynamics of the system which properties are estimated. In every time step $i$ a prediction $\hat{S}_i$ for the system state and its variance $\sigma_i^2$ can be predicted from the previously estimated state $\hat{S}_{i-1}$ and variance $\sigma_{i-1}^2$ by multiplying it with the matrix $A$ defining the transition of the system between two time points and considering process noise with the variance $\sigma_p^2$ corrupting the prediction:

$$S_i = A * \hat{S}_{i-1}, \sigma_i^2 = A^2 * \sigma_{i-1}^2 + \sigma_p^2$$  \hspace{1cm} (2)

The state estimate $\hat{S}_i$ and its variance $\sigma_i^2$ are then updated by integrating the prediction with the current measurement $y_i$ corrupted by measurement noise with the variance of $\sigma_y^2$ by weighting the prediction error with the Kalman gain:

$$K_i = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_y^2}$$  \hspace{1cm} (3)

*This work was supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 222641018 – SFB/TRR 135, A5.

1Anna Metzger and Knut Drewing are with the Department of General Psychology, Justus-Liebig University Giessen, 35394 Giessen, Germany anna.metzger@psychol.uni-giessen.de; knut.drewing@psychol.uni-giessen.de
\[ \hat{S}_i = S_i + K \ast (y_i - S_i), \hat{\sigma}_i^2 = \sigma_i^2 - K \ast \sigma_i^2 \] (4)

If the system is static the prediction of the Kalman filter model for the state of the system and the reliability of the estimate is identical with the MLE model (1). The Kalman filter was successfully used to model amongst others state estimation in visuomotor behavior (e.g., [16], [17] and lately also to approximate dynamic information integration in haptic perception within a single segment of exploration (i.e. one indentation in stiffness perception) [18]. By including memory loss in the optimal serial integration of information across different exploratory segments the weights of single estimates in haptic comparison of an object’s texture could be successfully predicted [13]. The memory loss was modeled by increasing the variance of the first stimulus’ estimate according to a power function of the number of strokes over the second stimulus with the exponent of 0.442. The exponent was taken from previous findings on memory decay in tactile two-point discrimination [19] and relating it to the change in the variance of perceptual estimates [13].

The model proposed by [13] considered only the case in which participants switch once between the stimuli and never switch back to the first stimulus, i.e. they explore the first stimulus than the second and then meet their decision. However, in a free haptic exploration, participants switch usually more often between the stimuli [20], [21]. And we could recently show that given the same amount of information (e.g. same number of indentations to compare the softness of two stimuli), more switches lead to better performance [22], suggesting that participants tune their exploratory behaviour to optimize performance. In the model of [13] the authors assume that after the exploration of the first stimulus a representation of its softness is built up. Once participants start to explore the second stimulus, in every indentation of the second stimulus a difference between the previous and the current stimulus is computed and integrated over time according to equations (2-4), considering that the representation of the first stimulus decays. The measurement noise of the difference \( \sigma_y \) in (3) was computed as the sum of the variance of the first stimulus and the variance of a single perceptual estimation. This assumes that the information about the currently explored stimulus is not integrated, inconsistent with the idea that information was integrated about the previous stimulus. Considering multiple switches between the stimuli estimation of their difference necessitates an additional representation and updating also of the currently explored stimulus, which is missing in the model of [13]. This is also more plausible given neurophysiological findings [23], [24]. In a task in which monkeys had to distinguish two subsequently presented vibrotactile stimuli \( f_1 \) and \( f_2 \), in the ventral premotor cortex (an area associated with sensorimotor tasks), during the presentation of the second vibration neurons were found which reflect the memory of \( f_1 \) encode the decisions \( f_1 > f_2 \) and \( f_1 < f_2 \), but also neurons encoding \( f_2 \) [23].

Here we provide a more general and more consistent model of the integration of sensory information in haptic perceptual comparisons by considering also switches between the stimuli and including representations of the two stimuli. We test the model by correlating empirically estimated performance in haptic softness discrimination and the predicted one. We take behavioral data from free softness explorations containing large variety of exploration patterns [20] and constrained softness explorations in which the number of switches was experimentally varied [22]. Additionally we assess whether natural haptic explorations are tuned to optimize performance by comparing the theoretically possible performance of different exploratory patterns and the frequency with which participants used those. We empirically showed that consistent with the prediction of our model due to memory limitations, for the same amount of available indentations an exploration involving more switches leads to better performance. Here we tested whether participants would prefer this exploration pattern in a free exploration.

II. METHODS

A. Model

We assumed that serially acquired information about the two stimuli and their difference is integrated optimally achieving the highest possible reliability according to equations (2-4). We estimated the standard deviation \( \sigma_y \) in (3) of the measurement noise (i.e. the precision in estimating the softness of a single stimulus) for every reference used in the behavioral experiments (see section B) from the average Weber fraction (17%) in softness perception estimated in [25] and dividing it by \( \sqrt{2} \) [26] to account for the fact that the Weber fraction was estimated from a comparison between two stimuli. The state transition matrix \( A \) was set to 1 because the softness of the stimulus is a static property and process noise \( \sigma_p^2 \) in (2) was assumed to be negligible. In our model during the exploration of the first stimulus information from every indentation \( i \) would be integrated with the previously obtained information to form an estimate \( \hat{S}_{1i} \) of the first stimulus softness with the variance \( \hat{\sigma}_{1i}^2 \). Once participants switch to the second stimulus in every indentation \( j \) of this stimulus they integrate information to form an estimate of its softness \( \hat{S}_{2j} \) with the variance \( \hat{\sigma}_{2j}^2 \). From the time point of the first switch also a difference \( D_k \) between the two stimuli can be estimated and integrated over subsequent indentations \( k \). Crucially, once a stimulus is not currently explored its representation in memory decays. We modeled the memory decay as proposed by [13] with an increase of the variance of the currently not explored stimulus according to a power function of the number of indentations of the current stimulus. Thus, in the \( j_{th} \) exploration of the second stimulus the variance of the first stimulus would increase according to \( \hat{\sigma}_{1j} = \hat{\sigma}_{1j} \ast \sqrt{0.442} \). In our model the variance of the difference estimate \( \hat{\sigma}_{Dk}^2 \) was then assumed to be the sum of the current variances of the softness estimates of the two stimuli. There were no free parameters in the model.
We identified single indentations by the time points in which we were not interested in the absolute values of the force. We were free to use as many indentations as they wanted on each stimulus and to switch back and forth between the stimuli as much as they wanted. There were four references and for each reference there were two comparisons one softer and the other harder. The compliances of the stimuli were the following: reference \(r_1\): 0.14 mm/N, comparisons \(c_1\): 0.12 mm/N, \(c_2\): 0.15 mm/N; \(r_2\): 0.21 mm/N, \(c_2\): 0.18 mm/N, \(c_3\): 0.24 mm/N, \(r_3\): 0.37 mm/N, \(c_3\): 0.29 mm/N, \(c_4\): 0.46 mm/N, \(r_4\): 0.74 mm/N, \(c_4\): 0.62 mm/N, \(c_5\): 0.88 mm/N. Every comparison and reference pairing was repeated for 192 times, totaling in overall 24576 trials.

We further took the behavioral data from [22] (later referred to as study 2). Here 11 participants completed a 2AFC task. They explored in every trial two silicon rubber stimuli by indenting them with the index finger of their dominant hand and had to report which one felt softer. In all trials each stimulus was overall indented four times but the number of switches was prescribed in every trial. Participants were allowed to switch between the stimuli either only once after four indentations of the first stimulus, three times after every second indentation or seven times. In the latter case they switched after every indentation. There was one reference (Young’s modulus of 59.16 kPa) and 9 comparison stimuli (Young’s moduli of 31.23, 42.84, 49.37, 55.14, 57.04, 69.62, 72.15, 73.29 and 88.18 kPa). For each switching condition the reference was paired with each comparison 12 times, resulting in overall 3564 trials.

Both studies were conducted at a visuo-haptic setup, in which vision of the stimuli and the exploring hand was occluded and position of the index finger of the dominant hand and the force exerted on the stimuli were recorded during exploration. For more details please see [20] and [22] respectively.

C. Data Analysis

From every trial in study 1 we extracted the exploration pattern, i.e. the number of indentations on each stimulus over time. For instance if a participant indented the first stimulus two times then switched to the other one and indented it three time, then went back to the first stimulus again and indented it once, the exploration pattern would be 2, 3, 1. From the exploration patterns we computed the length of the exploration as the sum of all the indentation on both stimuli and the number of switches. The exploration pattern was extracted from the force data. We first smoothed the force data with a moving average Gaussian window with a kernel of 360 ms similar to [20]. We have chosen a wider window than [20] for more robustness and because we were not interested in the absolute values of the force. We identified single indentations by the time points in which the derivative of the smoothed force data turned positive and the force exceeded 3 N. Similar to [20] we considered only indentations which were separated by intervals longer than 180 ms as different ones, to exclude movement pauses within an indentation. In study 2 the exploration pattern was prescribed by the condition, thus no additional analysis of the force was necessary.

Discrimination performance was computed in both studies as \(d'\). In study 1 we aggregated trials with the same exploration pattern and same reference stimulus across participants. Only exploration patterns were considered which occurred more than 100 times and had at least one wrong response, to ensure reliable estimates of performance. We computed \(d'\) as the inverse from the cumulative normal distribution at the average percent correct across participants and trials. In study 2 we aggregated trials with the same exploration pattern (i.e. same switching condition, see section B) and across comparison stimuli with a similar physical distance to the reference stimulus (stimuli pairs: 1 and 9, 2 and 8, 3 and 6, 4 and 5 see section B). In study 2 there were 264 trials for each exploration pattern and stimulus pair. We correlated the empirical performance with the one predicted by the model given the same exploration pattern, reference stimulus and difference between reference and comparison stimulus. Here \(d'\) was computed as \(\frac{\text{Preference}_{\text{comparison}} - \text{Preference}_{\text{reference}}}{\sigma_{\text{comparison}}}\) with \(n\) being the total number of computed difference estimates and \(\sigma_{d'n}\) being the standard deviation of the final difference estimate, i.e. the overall achieved discrimination precision. For statistical tests human performance and model predictions were normalized by computing the standard score for each reference and comparison pair to eliminate the effects of different discrimination difficulties on the correlation.

To assess whether natural haptic explorations are tuned to optimize performance we computed for each participant, each length of exploration consisting of \(Nr_{\text{indentations}} \in [3...8]\) the relative frequency with which each number of switches \(Nr_{\text{switches}} \in [1...Nr_{\text{indentations}} - 1]\) was used. We entered relative frequencies of usage into an one-way repeated measures ANOVA with the within-subjects factor \(Nr_{\text{switches}}\) to test whether some exploration pattern was preferred over others. Additionally we performed linear regression of the relative frequencies of usage on the number of switches to test our directional hypothesis that given the same length of the exploration more switches are preferred (i.e. the regression slope is significantly higher than 0). For the same exploration patterns as used by the participants we also computed the prediction of performance for the just noticeable stimulus difference using our model. Performance was averaged over explorations with the same length and number of switches (e.g. exploration patterns 1, 5 and 5, 1).

III. RESULTS

In study 1 from the total amount of data 61% was considered for analyses after applying the selection criteria. No data was excluded in study 2. Figure 1 shows the normalized observed performance as a function of the normalized
Fig. 1. A) Normalized observed discrimination performance in free explorations (study 1 [20]) as a function of normalized predicted performance given optimal integration of information with a Kalman filter and considering memory limitations. Brighter colors indicate longer explorations (i.e. consisting of more indentations) and larger symbols indicate explorations with more switches between the stimuli. B) Normalized observed discrimination performance in instructed explorations (study 2 [22]) as a function of normalized predicted performance. Here brighter colors indicate more number of switches. The length of the explorations was always the same. For both studies the observed and the predicted performance was normalized within trials with the same stimulus difference.

predicted performance for study 1 and 2. Normalization eliminates the effect of different reference and comparison distances on the correlation while preserving the order in performance given different exploration patterns. In study 1, as expected, exploration patterns with more indentations and more switches yield better performance as brighter colors coding longer explorations and bigger symbols coding for more switches occur more often in the upper range of observed and predicted performance. In study 2 the exploration length was the same in all conditions, but the increasing number of switches (increasing brightness) leads on average to better observed and predicted performance also here. In both studies we find a significant correlation between the observed and predicted performance, with correlation coefficients of $R = 0.66$, $p < 0.001$ and $R = 0.67$, $p = 0.018$, for study 1 and 2 respectively, indicating that the model overall successfully could predict the order of performance given different exploration patterns. However,
with normalization the relationship between the magnitudes of observed and predicted performance is lost, from which we could understand whether the model can also correctly predict the magnitude of performance. Figure 2 shows the not normalized data. In study 2 the performance is well predicted by the model in the lower range but in the higher range performance seems to saturate (i.e. the model overestimates the performance in the higher range). This is even more pronounced in study 2 where participants were instructed to indent each stimulus 4 times leading to a theoretically relatively high performance in all conditions.

Figure 3A shows for each length of exploration between 3 and 8 indentations in study 1 the average relative frequency of usage as a function of the number of switches. For this analysis 81% of data was considered after applying the selection criteria. For all exploration lengths there seem to be a preference for more switches up to around 3 switches. In longer explorations however, the frequency of usage falls off after 3 switches and raises again at the maximum available number of switches. The one-way repeated measures ANOVA with the within-subjects factor Nr_switches on relative frequencies of usage revealed for almost all exploration lengths a significant main effect of Nr_switches, or a trend. For the shortest and longest explorations the effect of Nr_switches on relative frequencies of usage was significant: 3 indentations, \( F(1,15) = 23.25, p < 0.001 \); 7 indentations, \( F(5,75) = 2.98, p = 0.017 \); 8 indentations, \( F(6,90) = 3.90, p = 0.002 \). There was a trend for 4 indentations, \( F(2,30) = 2.76, p = 0.08 \); and 6 indentations, \( F(4,60) = 2.42, p = 0.058 \). The effect was not significant for 5 indentations, \( F(3,45) = 1.78, p = 0.165 \). The slope in the regression of the number of switches was significantly positive or there was a trend only for the rather short explorations consisting of 3 indentations, \( t(15) = 4.82, p < 0.001 \); 4 indentations, \( t(15) = 1.68, p = 0.057 \) and 5 indentations, \( t(15) = 1.51, p = 0.075 \) (one-sided \( t \)-tests). However, it was significantly positive for all exploration lengths longer than 4 indentations in the range between 1 and 3 switches: 5 indentations, \( t(15) = 1.79, p = 0.047 \); 6 indentations, \( t(15) = 3.06, p = 0.004 \). 7 indentations, \( t(15) = 5.21, p < 0.001 \); 8 indentations, \( t(15) = 4.75, p < 0.001 \). Figure 3B shows the absolute frequencies with which differently long explorations were used. Shorter explorations were more frequently used than longer ones. Figure 3C shows the predicted performance for the exploration patterns used by participants for each exploration length as a function of the number of switches. The increase in performance is almost linear until 3 indentations but after it starts to saturate.

IV. DISCUSSION

We provide here a model for optimal serial integration of information in natural haptic perception, where integration of information is limited by memory. The previously proposed model [13] was limited to a specific case, i.e. explorations in which participants switch only once between the stimuli. Additionally, it was assuming that no information about the secondly explored stimulus was integrated, which is inconsistent with previous literature [23], [24] and would make switching between the stimuli impossible, which however happens very often in free haptic explorations. Thus it could not be applied to predict performance in free explorations. The model proposed by us is more consistent and general and thus allows predicting participants’ performance in natural multi-segmented haptic explorations. We found significant correlations between observed and predicted performance in two studies [20], [22], without any fit of the model to the data, indicating that our model could successfully predict the order in performance given different exploration patterns. It overall could also predict well the magnitude of performance for naturally used exploration patterns [20], indicating that integration of information is optimal given memory...
limitations in this circumstances. However, the performance of participants in [22] where they were instructed to use relatively long explorations (8 indentations in total) and an prescribed switching pattern, was mostly overestimated by the model. This can be due to the fact that in the instructed exploration participants’ memory and attentional capacities were also loaded with task requirements, negatively affecting their discrimination performance. However, also for the few natural explorations which are predicted to yield rather high performance (d’ > 2.5) there seem to be some saturation. This could indicate that performance might be limited by additional noise, not considered in our model. However, the fact that performance is well approximated by optimal integration for naturally used explorations, suggests that humans tend to preferentially use explorations in which optimal integration of information is possible.

We also investigated whether participants optimize their exploration to achieve high performance. Our model of integration information predicts that longer explorations lead to better performance and given the same length of the exploration more switches between the stimuli yield better performance (Fig. 3C). However, our model takes not into account that additional indentations and switches need additional muscular effort. Thus in the optimization of exploration both the maximization of performance and the minimization of muscular effort needs to be considered. For instance, the fact that for the extension of exploration additional muscular effort is necessary, manifests itself in the participants’ preference for shorter explorations over longer ones (Fig. 3B). For explorations with the same length, performance can be additionally optimized by switching between the stimuli. We found for all analyzed exploration lengths (3 to 8 indentations) in the range between 1 and 3 switches that participants indeed preferred to use more switches. However, in longer explorations after 3 switches the frequency of usage decreases. As also additional switching causes muscular effort it trades off with the improvement of performance. Interestingly, the decrease of the frequency in usage after 3 indentations coincides with the beginning of saturation in performance with increasing number of switches (Fig. 3C). Thus the exploration strategy of participants can be considered optimally in the sense that they invest additional muscular costs only as long there is maximal gain in performance.

ACKNOWLEDGMENT

We would like to thank Matteo Toscani for helpful discussions and suggestions throughout the project.

REFERENCES