Cahiers de Psychologie Cognitive/
Current Psychology of Cognition
2002, 21 (4-5), 537-564

A Bayesian view of motor adaptation

Alexander T. Korenberg and Zoubin Ghahramani

University College London, U.K.

Abstract

The adaptability of the motor system to both visual and force perturbation has been clearly demonstrated by experiments on adaptation to prismatic shifts of the visual scene or by adaptation to force-fields produced by a robotic manipulandum. While these experiments demonstrated the remarkable adaptability of the motor system, in the real world the adaptive system has to take account of the inherent uncertainty of the environment it interacts with. Two sources of uncertainty have to be considered: the uncertainty associated with sensory feedback and the uncertainty arising from the fact that the environment can change. These different forms of uncertainty give rise to constraints on how sensory and contextual information must be integrated over time in order to optimally update an adaptive control system. We treat motor adaptation as a system identification problem of inferring an internal model of the environment from noisy observations, using the example of a simple linear system and extend this idea to include contextual information. Bayesian theory provides a framework in which such uncertain sources of information can be integrated optimally. We briefly discuss experimental results for the role of contextual information on adaptation to dynamic loads and review a number of recent experiments that explicitly address the issue of adaptation to noisy and time-varying environments.

Key words: motor adaptation, motor learning, Bayesian inference, nonstationary environment, context, Kalman filter.

Correspondence should be sent to Alexander T. Korenberg, Gatsby Computational Neuroscience Unit, University College London, 17 Queen Square, London, WC1N 3AR, U.K. (email: alex@gatsby.ucl.ac.uk)

INTRODUCTION

On a typical day, one could ride a bike to work, walk up the stairs to the office, write down a quick note, type on a keyboard, use a computer mouse and maybe play a game of squash in the evening. What all these activities have in common is that they can be performed without any conscious awareness of the complicated underlying problem of producing the correct motor output given sensory information and a representation of the goal of the movements. They all involve different relationships between sensory inputs and motor outputs, or sensory-motor mappings, related to the different objects and environments we interact with. To complicate the matter further, these mappings are not fixed but can vary in time either predictably or in a random fashion: during a game of squash, muscle properties change due to fatigue; or we might have to deal with gusty side winds which will certainly introduce a random component in our bike ride home. This poses the dual question of how the motor system can adapt our behavior to a seemingly infinite number of dynamical environments and how we deal with the inherently noisy and time-varying characteristics of our world.

Since von Helmholtz observed that humans could adapt to displacements of the visual scene induced by wearing spectacles fitted with a set of wedge prisms (Helmholtz, 1925), the adaptability of the human motor system has fascinated researchers. A great number of experiments have been performed using the prism paradigm (see Welch, 1986, for a review) and more recently using computer aided set-ups (Ghahramani, Wolpert, & Jordan, 1996; Vetter & Wolpert, 2000b), testing the adaptability of the visuo-motor map. The adaptability of the visuo-motor map is thought to be crucial to keep the external world aligned with internal coordinate frames (Bedford, 1999), and probably explains the ease with which we can perform tasks like using a computer mouse on our desk to control a cursor on the screen.

While visuomotor adaptation to visual rearrangements has been studied for over a century, most of the above examples of possible activities during a typical day illustrate the importance of adaptation to environments or objects with different *dynamical* properties (e.g., inertia). This form of sensori-motor adaptation has begun to be explored more recently using programmable robotic manipulandi to perturb the hand's trajectory while subjects reached to targets displayed on a computer screen. Many studies have investigated how such a novel dynamic system is learned using position, velocity and acceleration dependent perturbations (Flash & Gurevich, 1997; Sainburg, Ghez, & Kalakanis, 1999; Shadmehr & Mussa-Ivaldi, 1994). This adaptation has been explained by assuming that the motor system forms an internal model of the perturbation to compensate for the perturbations in a predictive, feed-forward fashion. More recently this picture has been complemented by demonstrating the important role of rapid feedback loops for adaptation to dynamic loads (Bhushan & Shadmehr, 1999; Wang, Dordevic, & Shadmehr, 2001).

Given that we interact with an extremely large number of different and, usually, noisy and time-varying dynamical environments, how can we learn, store and recall internal models for numerous dynamic situations? Some studies have investigated the role of time (Conditt & Mussa-Ivaldi, 1999) and of contextual cues (Gandolfo, Mussa-Ivaldi, & Bizzi, 1996) in the acquisition of internal models for multiple dynamic situations, but adaptation to non-static, time-varying environments has only been studied recently (Korenberg & Ghahramani, 2002; Scheidt, Dingwell, & Mussa-Ivaldi, 2001; Takahashi, Scheidt, & Reinkensmeyer, 2001) and lacks a theoretical framework.

In this work we will develop a probabilistically consistent framework that combines estimation of time varying parameters, reflecting changes in the properties of the environment, with the influence of contextual cues. The problem is approached by interpreting motor learning as a system identification problem, using information about past experience, as well as the context, to infer a controller that results in the desired performance. We will review results from dynamic perturbation studies that investigate the role of contextual from a Bayesian perspective and interpret the results of studies that investigate adaptation to noisy, time-varying environments in our framework.

BAYESIAN INFERENCE FOR MOTOR ADAPTATION

Motor adaptation and the delta rule

Abstractly, the role of the motor system is to produce motor commands that achieve a given goal. In this view, the motor system is described by a function, or internal model, F that maps sensory feedback s to an output o (e.g., a motor command) such that as a result the goal is achieved,

$$o_t = F(w,s_t), \tag{1}$$

where the subscripted index t indicates that both o and s are a function of time. O is the output of the adaptive controller, for example a force compensating for an external perturbation, and s is any sensory information or history thereof, efference copy or state of mind relevant to producing o. If we assume that the desired output o^* is known, the goal of adaptation can be reduced to minimizing a measure E of the difference between desired and actual output of the system by tuning the parameters w. Often the square error is a convenient choice for E:

$$E = \frac{1}{2} \sum_t (o_t^* - o_t)^2 \tag{2}$$

E can then be minimized by taking small steps in the direction of decreasing E using gradient descent, which results in the well-known delta rule for updating w:

$$w_{new} = w_{old} - \eta \frac{\partial E}{\partial w} = w_{old} - \eta \cdot (o_t^* - o_t) \cdot \frac{\partial F}{\partial w} \tag{3}$$

In the remainder of this paper we will assume that F is scalar and linear in w; in particular we chose:

$$o_t = w \cdot s_t \tag{4}$$

This emphasizes the qualitative conclusions we would like to point out and we will mention the extension to the general n-dimensional and non-linear case in the discussion. Although this is an extremely simplified model of adaptation, it can capture the basic traits of simple adaptation experiments, for example, state based adaptation to a viscous force field that depends on the hand's velocity (Thoroughman & Shadmehr, 2000). In this example o^* corresponds to the force required to compensate for the force field, o is the actual force output of an adaptive controller implemented by the motor system and the relevant sensory variable s corresponds to the velocity of the hand. The assumptions of linearity lead to the simple update rule:

$$w_{new} = w_{old} + \eta \cdot (o_t^* - o_t) \cdot s_t \tag{5}$$

This form of the delta rule makes intuitive sense for changing the parameters w of the adaptive system, since it correlates the sensory input with the resulting error in the output. This gives it a biologically plausible interpretation as Hebbian learning between sensory neurons and neurons encoding an error signal. The learning rate η in this description is a free parameter that has to be set by hand, which is fine as long as we are only interested in reproducing a certain experimentally found adaptive behavior. There are however some conceptual issues with this approach: First the learning rate η has the wrong units for a learning rate, which in our definition of the update rule should be dimensionless; instead η has to be in units of the squared inverse of the units of s for Eq. (5) to be dimensionally consistent. Since η represents a parameter controlling neural plasticity, it is awkward for it to depend on details of how the peripheral signals are represented (in particular their magnitude). Second, when adapting to a non-stationary, time varying environment, there is a principled constraint on how to set η. Both these issues will be addresses in the next two sections.

Motor adaptation and system identification

In the previous section we discussed motor learning from a perspective of function approximation. The problem of inconsistent dimensions of the training signal is due to the choice of gradient descent as an optimization technique and there are alternative optimization methods, such as Newton's rule, which also have a nice biologically plausible and intuitive interpretation (Kawato, Furukawa, & Suzuki, 1987). On the other hand, the second problem of how to set the learning rate η in a time-varying environment is not addressed in the function approximation framework. To address this question, we will rephrase the motor learning problem in terms of system identification, meaning that we will look at the motor learner as a dynamical system described by an internal state variable (the adaptive coefficients w). In system identification, a model is identified by feeding a signal (e.g., white noise) to the system and observing the output. To apply this idea to motor learning, we assume that the adaptive system models the environment (in our simple example with a linear system) by identifying the state w that would result in the desired behavior, given the assumed model. In this view adaptation consists in inferring (identifying) an estimate w of the "true" coefficients w that would on average result in the desired output o_t^*. Given the simple output function in Eq. (4), and allowing for the fact that control signals are corrupted by noise (assumed to be Gaussian) the underlying model is:

$$o_t^* = w_t \cdot s_t + \delta_t; \quad \delta_t \sim N(0, \sigma_o^2) \tag{6}$$

where $\delta_t \sim N(0, \sigma_o^2)$ means that the noise δ_t is Gaussian distributed with mean 0 and variance σ_o^2. The desired output o_t^* is either available internally as a desired trajectory or directly from sensory input (i.e., the sensed external force which needs to be compensated), as is the sensory feedback s_t. We can therefore rewrite Eq. (6) in a more general form by defining $y = o_t^* / s_t$:

$$y_t = w_t + \delta_t; \quad \delta_t \sim N(0, \sigma_y^2) \tag{7}$$

We can see that in this form y is simply an observation of the state of the environment, corrupted by Gaussian noise. Although we started with a relationship between desired output and a model of the environment, Eq. (7) is a general observation model, capturing the fact that the motor system receives noisy sensory information about the environment.

Since we are interested in analyzing the more general case of a time-varying environment, we need to capture how the state of the environment w evolves with time. We again chose a very simple model assuming that the state of the environment drifts randomly (modeled by a random walk with Gaussian input noise). Although we choose this very simple model for simplicity, constantly drifting or decaying dynamics are equally easy to implement within the linear framework.

$$w_{t+1} = w_t + \chi_t; \quad \chi_t \sim N(0, \sigma_w^2) \tag{8}$$

where σ_w^2, the variance characterizing the random walk, models how fast the environment changes. A small value of σ_w^2 means that only small changes are likely to occur from one time step to the next and as a result w will drift slowly over time, while an environment characterized by a large σ_w^2 varies erratically from one time step to the next.

If the sensory observations would be noise free (i.e., $\delta = 0$ in Eq. (7)), then inferring w would be trivial, since $w = y$. However, in the presence of noise it becomes necessary to average sensory feedback y in order to obtain a reliable estimate of w. This can be done online as sensory information arrives:

$$\hat{w}_t = \hat{w}_{t-1} + K \cdot (y_t - \hat{w}_{t-1}) \tag{9}$$

where K is a dimensionless learning rate setting the effective number of observations being used in the averaging. The first observation about Eq. (9) is that unlike the delta rule, this update equation is dimensionally correct, and corresponds in fact to stochastic online version of Newton's method for optimizing w given the model of Eq. (7).

Setting K for inferring the state of a time-varying environment from noisy sensory feedback involves a trade-off between the amount of noise in the estimate, and the ability to track the changes in the environment. Setting K to a small value will lead to good noise-rejection, as each individual observation contributes only very little to the estimate, so that the estimator is effectively averaging over many observations (or equivalently a long time scale). On the downside, since for small K each observation has little effect on the estimate, a lot of observations are necessary to "convince" the estimator that a large change has occurred in the environment, or in other words the long time scale of averaging prevents the estimator from following rapid changes in the environment. Conversely a large value of K will result in an estimator that can respond to rapid changes, but will be much more variable. In summary the best setting of K should be large enough to track the changes in the environment, while rejecting as much noise as possible. This trade-off is

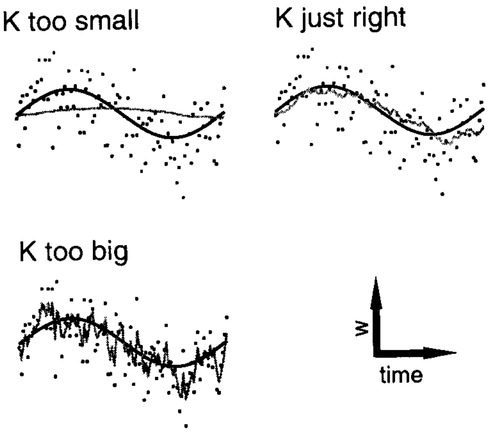

Figure 1. The trade-off between noise rejection and responsiveness.
When inferring a parameter w from noisy observations (dots) using online averaging over time t (Eq. (9)) a small learning rate K (top left panel) results in an estimate (gray line) with low noise but which does not follow the underlying "true" time course of w (black line). Conversely, if K is too large (bottom left panel), the estimate is very noisy. There exists an optimal setting of K (top right panel), that minimizes the estimation error by trading off noise rejection and responsiveness.

illustrated in Figure 1 which shows the behavior of an estimator described by Eq. (9) for an environment which has slowly varying state and noisy observations.

Can the intuition about this trade-off between consistency and responsiveness of the adaptive system be captured in a mathematical frame-work? In the next section we will discuss how this question can be addressed on a general level by considering the probability distributions over sensory feedback and the resulting estimate. For the case of a system with linear dynamics and Gaussian noise these considerations lead to the Kalman filter (Kalman, 1960), which implements online averaging (Eq. (9)) optimally. In addition to setting K in a principled way by taking account of the variability of the estimate and sensory feedback, the Kalman filter also incorporates the information from a predictive model (Eq. (8)) of the likely evolution of the state. This results in the optimal linear estimator in the sense that it results in the estimate with maximal posterior probability (MAP estimate), which also minimizes the squared estimation error.

Optimal inference and the Kalman filter

The dependence relationships implied by Eq. (7) and (8) can be compactly summarized in the graphical model shown in Figure 2. This graph represents the internal model the motor system is assumed to have built of the environment. The environment or object the motor system interacts with is represented by the hidden parameter w and causes the observed sensation y (see Eq. (7)) as indicated by the direction of the arrow. In addition the arrows between successive states w represent the dependence of w on previous time steps, given by Eq. (8).

The goal of the adaptive system is to calculate the best estimate of w_t given all previous sensory feedback $y_{t\ldots o}$ (i.e., all instances of sensory feedback y_t from the current time step $\tau = t$ back to the first time step $\tau = 0$). To find an optimal estimate the system needs to compute the posterior probability distribution of the state given all sensory feedback, $P(w_t \mid y_{t\ldots o})$, which can be calculated in terms of the distribution of the estimate given previous sensory feedback $P(w_t \mid y_{t\ldots o})$ (the prior over states), and the distribution of the current sensory feedback $P(y_t \mid w_t)$ (the likelihood of observing sensory feedback y given that the true state of the environment at time t is w_t) using Bayes rule (see Appendix):

$$P(w_t \mid y_{t\ldots0}) \propto P(y_t \mid w_t) \cdot P(w_t \mid y_{t\text{-}1\ldots0}) \tag{10}$$

The horizontal arrows in Figure 2 show us that the current value of w depends only on its value at the previous time step and the transition probability $P(w_t \mid w_{t\text{-}1})$. The predicted distribution of w based on previous feedback, $P(w_t \mid y_{t\text{-}1\ldots0})$, can therefore be found by averaging over the likely development of all possible values of the previous parameter $w_{t\text{-}1}$, weighted by their posterior probabilities:

$$P(w_t \mid y_{t\text{-}1\ldots0}) = \int P(w_t \mid w_{t\text{-}1}) \cdot P(w_{t\text{-}1} \mid y_{t\text{-}1\ldots0})\, \mathrm{d}w_{t\text{-}1} \tag{11}$$

Eq. (10) and (11) result in recursive formulas for updating the prior probability $P(w_t \mid y_{t\text{-}1\ldots0})$ to give the posterior $P(w_t \mid y_{t\ldots0})$ at each time step. Given the linear Gaussian assumptions we made, this recursive formula is known as the Kalman filter, and its derivation for our simple example system is developed in the Appendix.

The Kalman filter first updates the estimate and variance from the previous time step according to how the state of the environment is likely to have evolved given the assumed dynamics of w. Note that in our case only the variance is updated, since our model assumes that w

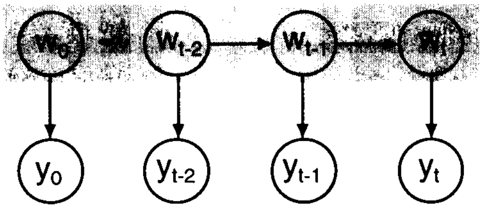

Figure 2. A graphical representation of the causal relationships implied by Eq. (7) and (8).
The state at time t, w_t, is assumed to drift over time, which is represented by the horizontal arrow linking states at subsequent time steps. The sensory observations y_t are assumed to depend on the state, as indicated by the vertical arrows.

follows a random walk, and there is hence no information on how w is likely to develop in the absence of sensory feedback:

$$\hat{w}_{t-\varepsilon} = \hat{w}_{t-1}; \quad \hat{\sigma}^2_{t-\varepsilon} = \hat{\sigma}^2_{t-1} + \sigma^2_w \tag{12}$$

where $\hat{w}_{t-\varepsilon}$ is the prior estimate, or prediction, before new sensory input becomes available and $\hat{\sigma}^2_{t-\varepsilon}$ is the associated uncertainty as measured by the variance of the of the prior distribution (see Appendix).

Second, the estimate and the associated uncertainty are updated to account for the newly arrived sensory feedback:

$$\hat{w}_t = \hat{w}_{t-\varepsilon} + K_t\left(y_t - \hat{w}_{t-\varepsilon}\right) \tag{13}$$

$$\hat{\sigma}^2_t = \hat{\sigma}^2_{t-\varepsilon} \cdot (1 - K_t) \tag{14}$$

The Kalman gain K_t defined according to Eq. (20) in the Appendix, determines how much an individual instances of sensory feedback influence the estimate of the underlying w:

$$K_t \equiv \frac{\sigma^2_{t-\varepsilon}}{\sigma^2_y + \sigma^2_{t-\varepsilon}} \tag{15}$$

These equations also make sense intuitively. The updating step in Eq. (12) always increases the uncertainty in the estimate, which is reasonable, since w has evolved randomly since the last observation. On the other hand, observing any sensory data (even with a very large variance) will always reduce the variance of the estimate (see Eq. (14)), and conversely, the variance of the estimate can never be larger than the variance of the sensory feedback (see Eq. (19)). The Kalman gain determines how much an individual instance of feedback influences the estimate, and hence sets the time scale over which sensory inputs are averaged. If the sensory feedback is very noisy relative to the uncertainty in the estimate, the Kalman filter averages over many time steps, reducing the noise. If on the other hand the feedback is relatively reliable, then larger changes to the estimate are made at each time step, resulting in a faster response of the estimator to changes in the environment. If no sensory feedback is available (equivalent to setting $\sigma_y{}^2 \to \infty$) the Kalman gain becomes zero and the only effect of the recursive update is to increase the uncertainty in the estimate due to the drift in the state of the environment.

The above procedure for optimal estimation is a special case of the Kalman filter for our assumption of a very simple dynamical system for w. In general the Kalman filter is applicable to any linear system (with arbitrary number of state variables) with Gaussian noise, and its derivation is along the same lines as presented here, using vectors and matrices to represent the state variables and their dynamics (Goodwin & Sin, 1984). For the cases of non-linear dynamics the system can be linearized around the previous estimate, resulting in the extended Kalman filter that, although not guaranteed to perform optimally, is generally a good approximation for systems with benign non-linearities. In an interesting example, Singhal and Wu (1989) have applied the extended Kalman filter to learning non-linear problems using a multi-layer neural network; this resulted in superior learning performance when compared to standard back-propagation algorithms. In this setting the Kalman filter can be seen as a way of finding the optimal learning rates for the neural network.

Although the Kalman filter uses the simplifying assumptions of linear dynamics and Gaussian noise, the principle of recursively applying Bayes rule (Eq. (10) and (11)) to optimally infer properties of the environment applies to arbitrary systems. It is however important to note that any optimal estimation procedure can only be said to be optimal given the knowledge of - and the assumptions about the system in question. As a result, it is crucially important to choose an appropriate model for the desired level of description.

Integrating contextual and temporal information

In the previous section we have derived the Kalman filter for a simple system underlying adaptation to a randomly changing, noisy environment. However, daily experience teaches us that we do not solely integrate past experience to determine our motor actions. Much to the contrary, in order to be able to interact with many different objects and environments, we must make efficient use of contextual information other than sensory feedback directly related to our action. When stepping on an escalator, for example, the context of seeing the escalator allows us to be prepared to step onto it. In fact, if the escalator is actually not moving, we experience a sensation of "surprise" for a short moment (Reynolds & Bronstein, 2001).

Fortunately, the Bayesian framework we used to derive the Kalman filter is easily extended to include context as an additional source of information. Consider the graphical model represented in Figure 3. The arrows indicate causal relationships between the variables, and the grey shading indicates that again w is a hidden parameter we want to estimate. In addition to sensory feedback y, we now have additional contextual information C helping us to infer the distribution of w. The graph can be summarized to mean that the object or environment associated with the hidden parameter w causes the sensation y, while the setting of the parameter depends on the context C. In the previous section we showed how Bayes rule can be used to infer an estimate of the state of the environment from information contained in the entire history of sensory feedback. Now we assume that we have access to contextual cues C, which give us direct information on the hidden parameter w. Using the fact that under this model sensory feedback y_t depends solely on the current hidden state w_t and that w_t depends only on past sensory feedback and context, we get:

$$\begin{aligned} P(w_t \mid y_{t\ldots0}, C_{t\ldots0}) &\propto P(y_t \mid w_t, y_{t-1\ldots0}, C_{t\ldots0}) \cdot P(w_{t-1} \mid y_{t-1\ldots0}, C_{t\ldots0}) \\ &= P(y_t \mid w_t) \cdot P(w_t \mid y_{t-1\ldots0}, C_{t\ldots0}) \end{aligned}$$

Unless we have a model of how the context and previous sensory experience interact to determine the hidden parameter, we cannot simplify this expression further. Fortunately it is often a good assumption that the influences of context and previous experience are independent. Consider, for example, the situation where the context indicates directly the nature of an object (and hence the likely value of w) by the feel of

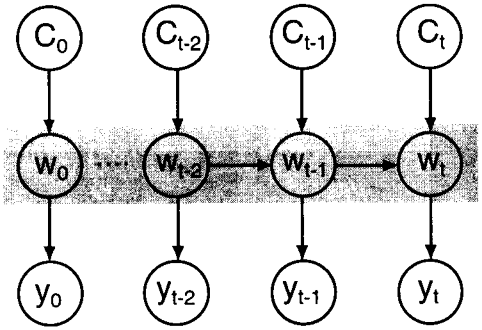

Figure 3. An extension to the model in Figure 2, where context affects the state at each time step. In this view the context determines a distribution over possible states, which is combined with sensory feedback online.

holding it in the hand. If we now pick up many such objects, the contextual cue indicates the nature of the object, independent of which object was picked up before. A possible counterexample in which context and previous experience interact would be a somewhat abstract cue indicating that the object will not change for some time.

Assuming that the influences of current context and past experience are indeed conditionally independent, i.e.,

$$P(w_t \mid y_{t-1\ldots0}, C_{t\ldots0}) \propto P(w_t \mid C_t) \cdot P(w_t \mid y_{t-1\ldots0}, C_{t-1\ldots0})$$

we get the posterior probability in terms of three factors:

$$P\left(w_t \middle| y_{t\ldots0}, C_{t\ldots0}\right) \propto \underbrace{\left[P(w_t \mid C_t)\right]}_{\substack{\text{current}\\ \text{context}}} \cdot \underbrace{\left[P(y_t \mid w_t)\right]}_{\substack{\text{current sensory}\\ \text{feedback}}} \cdot \underbrace{\left[P(w_t \mid y_{t-1\ldots0}, C_{t-1\ldots0})\right]}_{\substack{\text{integrated past}\\ \text{experience}}} \quad (16)$$

This expression for the posterior probability of w_t can be interpreted by picturing that both the current context and the integrated past experience are trying to wield their influence over the sensory feedback which had just been received. If, for simplicity, we assume that the distributions P are single-peaked (such as a Gaussian), then if the context is more informative of the setting of w, the distribution of w given the context (the first factor in Eq. (16)) will be narrower than the distribution given the previous experience (the last factor in Eq. (16)). When the distribution of the current sensory feedback (the middle factor in Eq. (16)) is multiplied by the two other factors, the location of its peak is more strongly attracted by the higher, narrower peak, as illustrated by Figure 4.

This way of combining different sources of information can explain why some form of contextual information can be very effective while others are ignored (see next section): if the motor system does not have an informative model $P(w_t \mid C_t)$ of how the context affects relevant parameters, i.e., $P(w_t \mid C_t)$ is flat and not particularly peaked on any value of w, then the context will have no or only very little influence on the final estimate of w. For example, the color of an object would not ordinarily be an informative context for the object's mass, whereas the size would be.

While the Gaussian assumption is plausible for the noise sources modeled by the Kalman filter, when considering contextual information, it might often be necessary to allow a multi-modal distribution of states given the context ($P(w_t \mid C_t)$ in Eq. (16)), for example, to represent a context in which we interact with one of two objects. Figure 5a-c illustrates that when the posterior distribution over states is calculated by combining a unimodal distribution (e.g., an estimate of the state based purely on sensory feedback) with a bimodal distribution representing the context, the result is not necessarily symmetric with respect to the MAP estimate. In this case the MAP estimate does not coincide with the least squares estimate (black and gray dots in Fig. 5).

In designing an estimator for this kind of bimodal problem, one has to consider which method to use to map the knowledge of the statistics of the world (contained in the posterior distribution over states) to an estimate which can be used to execute an action. This amounts to selecting one of two strategies: using the MAP estimate in the example illustrated in Figure 5c is equivalent to a switching strategy, since the estimate will always be jumping to the higher peak of the bimodal posterior. On the other hand, using the least square estimate is equivalent to

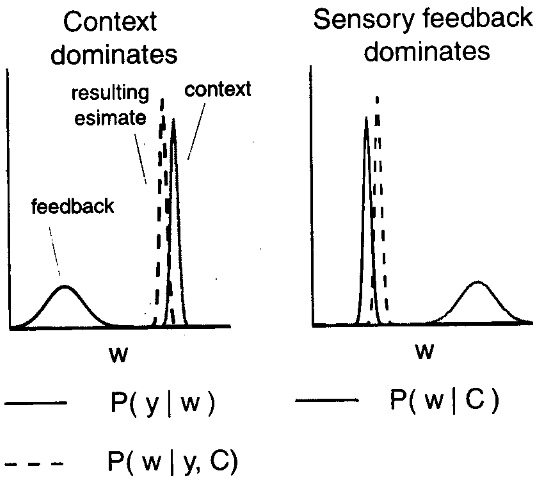

Figure 4. An example of how sensory feedback and contextual information are combined.
If the uncertainty associated with sensory feedback is large compared to the contextual information (i.e., $P(y|w)$ is broader than $P(w|C)$; left panel) then the resulting estimated is dominated by the contextual information, and vice versa (right panel). The graphs also demonstrate that the distribution of the resulting estimate is always narrower than each of the contributing sources of information, or in other words, adding even vague information reduces the overall uncertainty.

a weighting strategy corresponding to the full Bayesian treatment that minimizes a loss function (here the squared error) under the posterior distribution over states.

THE ROLE OF EXTERNAL CONTEXT FOR THE CONTROL OF MULTIPLE DYNAMICAL SETTINGS

When subjects learn to make reaching movements using a robotic manipulandum that produces velocity dependent force field, the movements are initially curved, but recover their initial shape after some practice (Shadmehr & Mussa-Ivaldi, 1994). If subjects subsequently are asked to make reaching movement in a force field which is the opposite

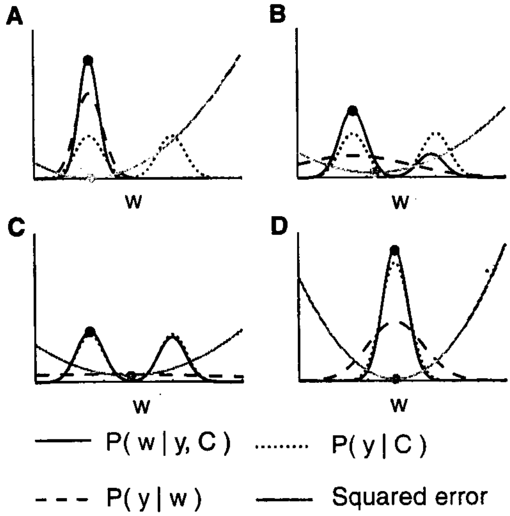

Figure 5. MAP and least square estimate do not necessarily coincide.
A-C: The distribution over sensory feedback (dashed line) is combined with a bimodal distribution containing contextual information (dotted lines) to give a probability distribution over *w* given sensory feedback and the context (solid lines). For progressively broader distributions over sensory feedback from A to C, the resulting combined distribution becomes more bimodal, and as a consequence MAP (black dot) and least square (gray dot) estimates move apart. The gray line shows the expected squared estimation error. D: If all sources of information are Gaussian, MAP and least square estimate always coincide.

of the first one, interference occurs between the two force fields (Brashers-Krug, Shadmehr, & Bizzi, 1996; Shadmehr & Holcomb, 1997): the learning of the second field is impaired, as compared to the initial acquisition of the first one, and learning the second field infers with the later recall of the first one. However, if a sufficient amount of time passes between the presentations of the two force fields (5-6 hours), there is no interference effect and the memory of the first field is spared. In other words, the controller for the first field some is somehow consolidated in memory. Moreover, this consolidation process can be correlated with metabolic changes in the brain (Nezafat, Shadmehr & Holcomb, 2001; Shadmehr & Holcomb, 1999; Shadmehr & Holcomb, 1997). In our framework this might be interpreted to suggest that learning of an estimate of the perturbation from sensory feedback is taking place in working memory, while storing the association between a given context and the perturbation $P(w|C)$ is a slower process requiring a period of consolidation. The functional imaging results support the idea that these two different sources of information might be stored in different parts of the brain. In this view, interference occurs because the online estimation process simply averages over past experience, while consolidation allows knowledge of the perturbation to be stored as a context dependent mapping from sensations (e.g., a force pushing to the right) to the correct setting of the controller.

The results of the above experiments show that multiple dynamic objects can be learned, if they are well separated in time and this raises the question what other contextual cues can be used to store controllers for multiple dynamic environment. While it seems obvious that in daily life the use of contextual information is crucial for successful interaction with our environment; the above observations raise the question of which kind of contextual information can be meaningful to the motor system. Gandolfo and colleagues (Gandolfo et al., 1996) observed that two opposed force-fields could be learned without interference if a different posture of the arm was used for each force field. After training, movements in the absence of the force fields resulted in posture specific aftereffects, indicating that a different controller was associated with each posture. However, this was only true for a postural change that involved joints participating in the movement - changing the posture of the thumb gripping the manipulandum did not have this effect. Similarly no effect was observed when the color of the room lighting was correlated with the force field.

These results can be described parsimoniously by postulating that adaptation is to some degree localized in state space (joint angles or workspace position and velocity). This interpretation implies that the motor system was unable to use the kind of cues presented in the above studies for the control of dynamic objects and this seems to be in conflict with our daily experience, where we interact with many objects based on visual, haptic, and other sensory feedback. However, in real life, the motor system might rely on a coincidence of many consistent sensory instances to define an object or environment, e.g., visual appearance, as well as the "feel" of a tennis racket, so that the simple kind of contextual information provided in the experiments is simply too uninformative. In the formalism we developed above, this correspond to a flat probability distribution $P(w|C)$ over adaptive states given the context of, e.g., room lighting. In fact it would be rather surprising to find any strong effects of such artificial contexts - the motor system has a whole lifetime of exposure to naturally occurring contextual information, so that it is not surprising that arbitrary associations, such as the association between the color of room lighting and the dynamical properties of a manipulated object, should be hard to learn.

CONTROL OF MOVEMENTS IN TIME-VARYING ENVIRONMENTS

The Bayesian framework put forward in this paper tells us how feedback should be integrated over time to optimally infer a controller for noisy and time-varying environments or objects. Control in such environments has been studied experimentally for both "prism adaptation paradigms" involving kinematic perturbations of visual feedback, as well as force-field adaptation paradigms which perturb the dynamic properties of a manipulated object.

Adaptation to kinematic perturbations

How does the internal estimate of a visual perturbation depend on the availability of visual feedback? To address this question, Vetter and Wolpert (2000b) asked subjects to make reaching movements to a visual target. After an initial learning phase, when subjects experienced a random sequence of two visual perturbations, they were tested with either no visual feedback or brief instances of feedback. Their results can be summarized as showing that in the absence of visual feedback, the motor system had a prior estimate of the perturbation given the context of the experimental setting, which was the average of the two perturbations in the learning phase. Incidentally, this suggests that the motor system averages over all possible context instead of using a MAP estimate, as discussed above. Furthermore, the effect of feedback on the estimate of the perturbation, as measured by the position of the endpoint of the movement was larger if feedback was given later on in the movement. Although this result is seemingly paradoxical, since receiving feedback earlier leaves more time for the movement to be updated, it is naturally explained by our estimation framework, which assumes that sources of information are weighted by a measure of their certainty: in the absence of sensory feedback, the uncertainty in the feedback-based estimate grows with time according to Eq. (8) and (12); consequently estimates based on later instance of feedback will have lower uncertainty at the end of the movement and will hence be weighted more strongly.

The results in the previous paragraph suggest that internal estimates of the environment are updated, even in the absence of sensory feedback, as suggested by Eq. (8). Another experiment in the same paper illustrates this point further: if subjects make reaching movements with either veridical visual feedback or feedback translated relative to the hand position, then the hand position does not decay to baseline when feedback is removed for a number of reaches. Instead it converges to the average of the veridical and perturbed hand positions. Again this can be explained by the increasing uncertainty in the absence of the visual feedback, so that the context based information (equal amounts of veridical and translated trials) dominates. Vetter and Wolpert model this prior knowledge implicitly by assuming a process switching between two discrete, modules (one for each context), each with a prior probability of 0.5. However, the data are equally well explained by a non-modular model of sensory and contextual integration of the form of Eq. (16): as feedback is removed, the uncertainty associated with the feedback based estimate increases, so that the final estimate is eventually dominated by the contextual information. The context could either be represented as a Gaussian centered on the average translation, or a bimodal distribution with a lobe at both the veridical and translated settings (see Fig. 5c,d).

Most of the models in this paper assume that the CNS has a very simple random walk model (Eq. (8)) of how the world evolves. This raises two questions: First, in an environment which is well described by this model, how close to the optimal estimator (the Kalman filter given by Eq. (13), (14), and (15)) is human performance? Ingram (1999) shows that human subjects perform very close to the ideal estimator in a reaching task with randomly drifting displacements of the visual feedback of the endpoint. She found that subjects were able to use up to 80 percent of the information contained in the sequence of displacements. Second, is the motor system limited to a random walk model (Eq. (8)) in estimating the evolution of a given context? Using a tracking task, Vetter and Wolpert (2000a) provide evidence that the CNS can at least model a constant rate of change when exposed to a slow, sinusoidally varying visual perturbation.

Adaptation to dynamic perturbations

The visual perturbation experiments discussed in the previous section rely on the fact that for a purely visual task, sensory feedback can easily be manipulated. Interpretation of these experiments relies on the idea that, since there is no mechanical interaction with the movement, the endpoint of the movement is a good indicator of the internal estimate, or model, of the perturbation. Wolpert and colleagues used an elegant "trick" to extend this idea to dynamic perturbations (Wolpert, Ghahramani, & Jordan, 1995) by asking the subjects to estimate the position of their hand after making a movement in the dark. A Kalman filter model similar to the one described in this paper accurately predicted the changes in the bias and variance of the estimate of hand position, both as a function of movement direction and externally applied forces.

However, here we are not interested in how the motor system forms an internal estimate of its own state, but the related question of how an internal estimate of the dynamics of an environment or object is formed and used for control. In that case it is impossible (at least in healthy subjects) to directly dissociate the mechanical perturbations applied to the hand from the feedback received by the motor system, as proprioceptive sensory feedback will always be present. While catch trials, where the external perturbation is unexpectedly removed, can give an indication of the internal estimate the motor system has formed about the environment, the interpretation can be confounded by the continuous presence of feedback. Nevertheless recent experiments on non-stationary force fields have provided some results relevant to the proposed estimation problem. We will discuss these findings in the light of a system identification perspective.

When subjects made reaching movements while exposed to a velocity dependent force-field with a magnitude that varied randomly from one trial to the next, the effect of the perturbation on the movements decreased during the experiment (Scheidt et al., 2001; Takahashi et al., 2001). This suggests that subjects were able to adapt to the force fields, in spite of their random nature. Both studies found that subjects showed an adaptive state appropriate for the average magnitude of the force field, even when the force-field magnitude had a bimodal distribution and the average magnitude field was never experienced (Scheidt et al., 2001). Scheidt and colleagues also found a dependence of performance on the force field experienced on the previous trials, indicating that a estimate of the force field is based on previous experience, as previously shown for a single, fixed force field (Thoroughman & Shadmehr, 2000). Unfortunately, both these results are consistent with *any* learning rule based on incremental updates as new sensory data becomes available, which would result in the reported averaging characteristics, as well as the dependence on previous experience.

However, Takahashi and colleagues (2001) observed that while catch trials were curved in the direction opposite to the force field (indicating that an internal estimate of the force was indeed formed), this curvature was smaller for catch trials following exposure to the random field than for catch trials following a fixed force field with a magnitude equal to the average magnitude of the randomly varying one. They interpreted this finding as evidence for a dual control mechanism: an estimation based controller compensating for the average perturbation, in combination with low-level impedance control by regulating muscle-cocontraction. While stiffness regulation surely forms an important part of the control strategies available to the human motor system, these experimental findings are parsimoniously explained by the Kalman filter model: for the fixed force field, the Kalman gain will be small because the estimate based on previous experience is very reliable. In the more variable, random condition, the Kalman gain would be increased because of the increased uncertainty associated with the estimate, giving higher weight to online feedback. As a result, the absence of the force on catch trials is detected earlier in the trial and the resulting movement is less curved. Although both the co-contraction and the Kalman filter hypothesis can explain the decreased magnitude of the catch trials, the experiment discussed in the next paragraph supports the latter.

The studies discussed in the previous section show that in the case of adaptation to a randomly switching force field, movements are consistent with adaptation to the average of the experienced forces. This suggests that in the case when the average perturbation is zero no adaptation occurs. This idea was tested by Korenberg and Ghahramani (2002), by investigating reaching movements in a force field switching between two opposed directions. They studied both a random sequence of the two force fields, as well as a predictable sequence, switching the force on every trial. All but two (out of a total of 28) subjects were unable to use the information contained in the predictable sequence to compensate for the force-field, which was evident from the fact that catch trials where only slightly curved and that there was no significant difference between random and predictable sequences. Furthermore, the small curvature of catch trials was opposite to the direction of the previously experienced (as opposed to one predicted by the sequence), indicating a residual adaptation to the previously experienced force field. This indicates that in nearly all of the subjects, adaptation was purely based on previous experience, even if the perturbation was completely predictable from the preceding trial.

In spite of the absence of feedforward adaptation, the reaching movements became significantly straighter during the experiment and, consistent with the results by Takahashi and colleagues (2001), the bias of catch trials decreased. However, Korenberg and Ghahramani (2002) found additional changes to the pattern of reaching movements, that support the hypothesis that the CNS indeed integrates sensory feedback according to the estimation model in Eq. (13)-(15). The underlying assumption is that since sequence information is not used, the switching force field is interpreted as a rapidly drifting environment, and hence sensory feedback is given a greater weight. This makes the prediction that feedback should affect the ongoing movement progressively more, and that movement variability should increase, due to the reduced time scale of averaging.

To test the effect of feedback on the ongoing movement, Korenberg and Ghahramani introduced a novel type of catch trial for which the force was removed only after the movement started and found an increasing response to the resulting small force pulse at the beginning of the movement, as predicted by the model. This increasing feedback response suggests that increasing co-contraction is unlikely to explain the changes in reaching movements, since increased co-contraction could be expected to lead to higher limb impedance. In addition, the variability of catch trials was increased more than 10 fold with respect to baseline movements although both types of movements have exactly the same dynamic context (i.e., forces are switched off). This is consistent with the second prediction of the model that everything else being equal, the variability of movements should increase as a result of the exposure to a switching force field environment.

In summary, the switching force field experiments have two main conclusions. Consistent with the difficulty in associating a force field with an abstract cue (e.g., room lighting), the information contained in a simple, predictable sequence of opposed force fields could not be used to learn both fields simultaneously. Instead performance was not significantly different if a random sequence was used. The behavior was consistent with an adaptive mechanism based on online averaging of sensory information, as in the case of a fixed force field (Thoroughman & Shadmehr, 2000) or a force field with randomly varying magnitude, but fixed direction (Scheidt et al., 2001; Takahashi et al., 2001). However, in the case of two opposed force fields, adaptation to one field interferes with performance when moving in the other field. Under these circumstances, the influence of previous trials decreases, while the effect of feedback online increases. These findings are parsimoniously explained by the integration model proposed in this paper, assuming that the motor system approximates the switching force field environment by a rapidly drifting one.

CONCLUSION

We have presented a new point of view for the analysis of motor learning which explicitly takes into account the uncertainty inherent in the sensory feedback and the controlled object. The key idea is to treat motor learning as the problem of inferring the parameters of the control system from noisy observations and contextual information. In this framework Bayesian theory is used to optimally combine these different sources of information. In the case of a linear system with Gaussian noise, this approach results in the Kalman filter, which is the optimal estimator under these assumptions.

We based our discussion on an extremely simplified linear system, as our goal was to illustrate the basic idea of how uncertainty in sensory feedback *and* the controlled environment leads to constraints on the efficient use of these sources of information. Using such a simplified model necessarily means renouncing the ambition to explain many experimental results in great detail, for example, details of movement trajectories. In fact by limiting the complexity of the model, we explicitly ignore details of the motor system in order to concentrate on the underlying principle of optimal probabilistic inference as a unifying explanation for a number of experimental observations relating to adaptation in non-stationary environments. Rather than explaining the details of each of the experiments discussed in this paper, the aim of the model is to point out and explain the communalities. Such a review of experimental work in the light of an abstract model is necessarily incomplete, although we have included all experimental work relevant to the computational modelling of adaptation to non-stationary environment that we are aware of. Likewise, our model is not comprehensive of all possible reactions of the motor system in the face of non-stationary environments, as it completely ignores, for example, any high-level cognitive influences.

We used a very simple random walk model of how the environment might change over time. Considering that a large amount of the uncertainty arises from the motor system itself (e.g., neuro-motor noise, fatiguing muscles, miscalibration of visuomotor maps), modelling the controlled environment as randomly drifting might actually be a close approximation to the reality the motor system has to deal with. In fact, when subjects are exposed to a visual perturbation with such a drifting behavior, the resulting behavior is close to optimal in the sense that most of the information available in the random sequence of perturbations is used (Ingram, 1999).

On the other hand in the switching force-field experiment (Korenberg & Ghahramani, 2002), subjects acted far from optimal by completely discarding the information available from the predictable alternation of force-fields. The key to understanding this discrepancy might be to consider the ecological environment that the motor system has evolved to master. In order to be able to rapidly adapt its behavior to the environment based on limited, noisy data, the motor system has used prior knowledge of the statistical structure of the world. Real world objects simply do not switch in predictable sequences and without associated sensory cues; it therefore seems plausible that artificial environment of this kind could be interpreted to be drifting rapidly instead.

A classic experiment in animal conditioning of taste aversion underlines the importance of the animal's prior assumptions in how it interprets contextual information. If a rat was given a combined conditioned stimulus of sugar water, light and sound, and then made to feel sick, only the taste of the sugar water resulted in aversive behavior. However, if an electric shock was used as unconditioned stimulus, it became associated with the light and sound (Garcia & Koelling, 1966) and not with the sugar water. The associations the rat formed were determined by the set of naturally occurring associations - sickness being associated with taste (e.g., from ingesting poisoned food) and pain with an external event. Similarly, while the human animal readily uses many types of contextual cues to interact with the world, the force field adaptation studies discussed here have shown that non-ecological cues such as room lighting or predictable switching sequences are not easily associated with behaviors which are otherwise easily learned.

RÉSUMÉ

L'adaptabilité du système moteur à des perturbations visuelles ou mécaniques a été clairement démontrée par des expériences d'adaptation à des décalages prismatiques de la scène visuelle, ou par l'adaptation à des champs de force produits par un manipulandum robotisé. Bien que ces expériences aient démontré la remarquable adaptabilité du système moteur, dans le monde réel, le système adaptatif doit tenir compte de l'incertitude inhérente à l'environnement avec lequel il interagit. Deux sources d'incertitude doivent être considérées : l'incertitude associée aux retro-actions sensorielles et l'incertitude provenant du fait que l'environnement peut changer. Ces différentes sortes d'incertitude produisent des contraintes sur la façon dont les informations sensorielles et contextuelles doivent être intégrées dans le temps afin de mettre à jour de façon optimale un système de contrôle adaptatif. Nous considérons l'adaptation motrice comme un problème d'identification de système inférant un modèle interne de l'environnement à partir d'observations bruitées, en nous basant sur l'exemple d'un système linéaire simple, et en étendant cette idée afin d'inclure l'information contextuelle. La théorie Bayésienne fournit un cadre dans lequel de telles sources incertaines d'information peuvent être intégrées de façon optimale. Nous discutons brièvement les résultats expérimentaux en termes de rôle de l'information contextuelle sur l'adaptation à des charges dynamiques et nous passons en revue un certain nombre d'expériences récentes qui abordent explicitement le thème de l'adaptation à des environnements bruités et variant en fonction du temps.

REFERENCES

Bedford, F. L. (1999). Keeping perception accurate. *Trends in Cognitive Sciences, 3,* 4-11.

Bhushan, N., & Shadmehr, R. (1999). Computational nature of human adaptive control during learning of reaching movements in force fields. *Biological Cybernetics, 81,* 39-60.

Brashers-Krug, T., Shadmehr, R., & Bizzi, E. (1996). Consolidation in human motor memory. *Nature, 382,* 252-255.

Conditt, M. A., & Mussa-Ivaldi, F. A. (1999). Central representation of time during motor learning. *Proceedings of the National Academy of Sciences of the USA, 96,* 11625-11630.

Flash, T., & Gurevich, I. (1997). Models of motor adaptation and impedance control in human arm movements. In P. Morasso & V. Sanguineti (Eds.), *Self-organization, computational maps and motor control* (pp. 423-481): Amsterdam: Elsevier Science.

Gandolfo, F., Mussa-Ivaldi, F. A., & Bizzi, E. (1996). Motor learning by field approximation. *Proceedings of the National Academy of Sciences of the USA, 93,* 3843-3846.

Garcia, J., & Koelling, R. A. (1966). The relation of cue to consequence in avoidance learning. *Psychonomic Science, 4,* 123-124.

Ghahramani, Z., Wolpert, D. M., & Jordan, M. I. (1996). Generalization to local remappings of the visuomotor coordinate transformation. *Journal of Neuroscience, 16,* 7085-7096.

Goodwin, G. C., & Sin, K. S. (1984). *Adaptive filtering, prediction, and control.* Englewood Cliffs, NJ: Prentice-Hall.

Helmholtz, H. v. (1925). *Treatise on physiological optics* (1867). Rochester, New York: Optical Society of America.

Ingram, H. (1999). *Sensorimotor integration and control in human movement.* Unpublished DPhil, University of Oxford, Oxford.

Kalman, R. E. (1960). A new approach to linear filtering and prediction. *Journal of Basic Engineering, Trans ASME D, 82,* 35-45.

Kawato, M., Furukawa, K. & Suzuki, R. (1987). A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics, 57,* 169-185.

Korenberg, A. T., & Ghahramani, Z. (2002). Adaptive feedback control for non-stationary environments. Gatsby Computational Neuroscience Unit Technical Report.

Nezafat, R., Shadmehr, R., & Holcomb, H. H. (2001). Long-term adaptation to dynamics of reaching movements: a PET study. *Experimental Brain Research, 140,* 66-76.

Reynolds, R., & Bronstein, A. (2001). Aftereffect of walking onto a moving platform. Paper presented at the 3rd conference on sensorimotor control in men and machines, Marseille.

Sainburg, R. L., Ghez, C., & Kalakanis, D. (1999). Intersegmental dynamics are controlled by sequential anticipatory, error correction, and postural mechanisms. *Journal of Neurophysiology, 81,* 1045-1056.

Scheidt, R. A., Dingwell, J. B., & Mussa-Ivaldi, F. A. (2001). Learning to move amidst uncertainty. *Journal of Neurophysiology, 86,* 971-985.

Shadmehr, R., & Holcomb, H. H. (1997). Neural correlates of motor memory consolidation. *Science, 277,* 821-825.

Shadmehr, R., & Holcomb, H. (1999). Inhibitory control of competing motor memories. *Experimental Brain Research, 126,* 235-251.

Shadmehr, R., & Mussa-Ivaldi, F. A. (1994). Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience, 14,* 3208-3224.

Singhal, S., & Wu, L. (1989). Training multilayer perceptrons with the extended Kalman algorithm. Paper presented at the Advances in Neural Information Processing Systems.

Takahashi, C. D., Scheidt, R. A., & Reinkensmeyer, D. J. (2001). Impedance control and internal model formation when reaching in a randomly varying dynamical environment. *Journal of Neurophysiology, 86,* 1047-1051.

Thoroughman, K. A., & Shadmehr, R. (2000). Learning of action through adaptive combination of motor primitives. *Nature, 407,* 742-747.

Vetter, P., & Wolpert, D. M. (2000a). The CNS updates its context estimate in the absence of feedback. *Neuroreport, 11,* 3783-3786.

Vetter, P. & Wolpert, D. M. (2000b). Context estimation for sensorimotor control. *Journal of Neurophysiology, 84,* 1026-1034.

Wang, T., Dordevic, G. S., & Shadmehr, R. (2001). Learning dynamics of reaching movements results in the modification of arm impedance and long-latency perturbation response. *Biological Cybernetics*, in press.

Welch, R. B. (1986). Adaptation to space perception. In K. B. Boff, L. Kaufmann, & K. E. Browman (Eds.), *Handbook of perception and performance* (Vol. 1, pp. 24.21-24.45). New York: Wiley-Interscience.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science, 269*, 1880-1882.

APPENDIX - BAYES RULE

Consider two events A (e.g., "it rained") and B (e.g., "*the ground is wet*") occurring with probabilities $P(A)$ and $P(B)$. If the two events are not independent, then *given* that A occurred we get a different probability $P(B|A)$ for observing B. $P(B|A)$ (i.e., the probability that the ground is wet, given that it rained) is called the conditional probability of B given A, and similarly we can define the probability of $P(A|B)$ (i.e., the probability that it rained, given that the ground is wet). The joint probability of both A and B (i.e., wet ground and rain) occurring is given by:

$$P(A,B) = P(A)\cdot P(B \mid A) = P(B)\cdot P(A \mid B)$$

Let's suppose we observe B but are really interested in knowing how probable A is given that we observed B. To do this inference, we can use Bayes rule, which follows straight from the above expression for the joint probability:

$$P(A \mid B) = \frac{P(B \mid A)\cdot P(A)}{P(B)}$$

In this context, $P(B|A)$ is called the likelihood of observing B, $P(A)$ is the prior and $P(A|B)$ is the posterior probability over A. The same argument holds if the Ps are probability distributions over the random variables A and B. In that case $P(B)$ is a constant with respect to A, ensuring that $\int P(A \mid B)dA = 1$, as required for a probability distribution. If we are only interested in the shape of $P(A|B)$, than we do not need to be concerned with this normalization constant and can use

$$P(A \mid B) \propto P(B \mid A)\cdot P(A)$$

APPENDIX - A SIMPLE KALMAN FILTER

This appendix derives the Kalman filter for the optimal estimate given the assumptions of linear dynamics (in fact a random walk) and Gaussian noise. The derivation reads on from Eq. (11) and results in the update equations (12), (13) and (14). Given the assumptions of linear dynamics and Gaussian noise for the evolution of w (Eq. (8)), the prior distribution $P(w_t \mid y_{t-1\ldots0})$ is also a Gaussian: Since we assume that w follows a random walk with no consistent trend, $P(w_t \mid y_{t-1\ldots0})$ is also a Gaussian with the same mean and increased variance as compared to $P(w_{t-1} \mid y_{t-1\ldots0})$. To indicate that the prior is in fact the distribution representing our knowledge about w just before the current measurement is taken into account, we use the time index t-ε:

$$P(w_t \mid y_{t-1\ldots0}) \sim N(\mu_{t-\varepsilon},\sigma^2_{t-\varepsilon}); \qquad (17)$$

$$\mu_{t-\varepsilon} = \mu_{t-1}; \quad \sigma^2_{t-\varepsilon} = \sigma^2_{t-1} + \sigma^2_w$$

The sensory feedback is given by the state of the world corrupted by Gaussian noise and we therefore also get a Gaussian (centered on w and with variance given by the feedback noise) for the likelihood of observing a given sensation y:

$$P(y_t \mid w_t) - N(w_t, \sigma_y^2) \tag{18}$$

Since the posterior probability $P(w_t \mid y_{t...0})$, which we are interested in, is simply the product of two Gaussians (Eq. (17) and Eq. (18)), it is also a Gaussian, centered on the average of the individual centers weighted by the inverse of the respective variances and with a variance equal to the geometric mean of the individual variances:

$$P(w_t \mid y_{t...0}) \sim \mathcal{N}\left(\mu_t, \sigma_t^2\right)$$

$$\mu_t = \frac{1/\sigma_{t-\varepsilon}^2 \cdot \mu_{t-\varepsilon} + 1/\sigma_y^2 \cdot y_t}{1/\sigma_{t-\varepsilon}^2 + 1/\sigma_y^2}; \quad 1/\sigma_t^2 = 1/\sigma_{t-\varepsilon}^2 + 1/\sigma_y^2 \tag{19}$$

How do we extract an estimate of w_t from the posterior distribution? One possibility is to simply pick the value of w_t for which the posterior probability is maximized (MAP estimate), i.e., the mean of the distribution in the Gaussian case. Alternatively we could seek to minimize the estimation error. If we choose the squared estimation error as a cost function, it is easy to show that the resulting (least square) estimate is again the mean of the posterior distribution. Although in the Gaussian case (and in fact for any unimodal distribution centered on its mean) the MAP and least square estimates coincide, this is not generally the case.

By equating the estimate $\hat{w}$ with the mean μ of the posterior and rewriting Eq. (19) as

$$\mu_t = \mu_{t-\varepsilon} + \frac{\sigma_{t-\varepsilon}^2}{\sigma_y^2 + \sigma_{t-\varepsilon}^2}\left(\mu_{t-1} - y_t\right); \quad \sigma_t^2 = \left(1 - \frac{\sigma_{t-\varepsilon}^2}{\sigma_y^2 + \sigma_{t-\varepsilon}^2}\right)\sigma_{t-\varepsilon}^2 \tag{20}$$

we obtain a recursive procedure for calculating the optimal estimate $\hat{w}$ of w and the associated uncertainty, as measured by the variance in $\hat{w}$, $\hat{\sigma}^2$. Using the definition of the Kalman gain in Eq. (15) result in the update equations Eq. (12), (13), and (14).

Received 26 October, 2001
Accepted 23 May, 2002