Published in final edited form as:

J Neural Eng. 2005 September ; 2(3): S147–S163.

Internal Models in Sensorimotor Integration: Perspectives from Adaptive Control Theory

Chung Tin^{1,2} and Chi-Sang Poon²

¹ Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

² Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract

Internal model and adaptive control are empirical and mathematical paradigms that have evolved separately to describe learning control processes in brain systems and engineering systems, respectively. This paper presents a comprehensive appraisal of the correlation between these paradigms with a view to forging a unified theoretical framework that may benefit both disciplines. It is suggested that the classic equilibrium-point theory of impedance control of arm movement is analogous to continuous gain-scheduling or high-gain adaptive control within or across movement trials, respectively, and that the recently proposed inverse internal model is akin to adaptive sliding control originally for robotic manipulator applications. Modular internal models architecture for multiple motor tasks is a form of multi-model adaptive control. Stochastic methods such as generalized predictive control, reinforcement learning, Bayesian learning and Hebbian feedback covariance learning are reviewed and their possible relevance to motor control is discussed. Possible applicability of Luenberger observer and extended Kalman filter to state estimation problems such as sensorimotor prediction or the resolution of vestibular sensory ambiguity is also discussed. The important role played by vestibular system identification in postural control suggests an indirect adaptive control scheme whereby system states or parameters are explicitly estimated prior to the implementation of control. This interdisciplinary framework should facilitate the experimental elucidation of the mechanisms of internal model in sensorimotor systems and the reverse engineering of such neural mechanisms into novel brain-inspired adaptive control paradigms in future.

1. Introduction

The notion of internal model in sensorimotor integration has attracted increasing attention in recent years particularly in the areas of motor control and vestibular motion sensing (Kawato 1999; Merfeld, Zupan et al. 1999; Angelaki, Shaikh et al. 2004; Kording and Wolpert 2004). Internal model is a hypothetical central neural representation of the dynamics and kinematics of movement, which is believed to underlie the nervous system's remarkable ability to discern unknown or underdetermined changes in the environment. For example, forward and inverse internal models in motor control reconstruct respectively the direct and inverse dynamics of the motor system (Kawato 1999). The vestibular system estimates the ambiguous system states with an internal model in conformance with physical laws of mechanics (Merfeld and Zupan 2002). The study of internal models in these areas has shed new lights on the mechanisms of sensorimotor integration and learning in the brain.

Corresponding author: Chi-Sang Poon, Ph.D., Harvard-MIT Division of Health Sciences & Technology, Bldg. 56-046, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA, Tel: +1 617-258-5405; Fax: +1 617-258-7906, Email: cpoon@mit.edu.

While a novel concept in its own right, the internal model paradigm is in many respects reminiscent of the classic adaptive control theory in engineering. The study of adaptive control was first conceived in designing autopilots for high-performance aircrafts that operate over a wide range of speeds and altitudes. The dynamics of the aircraft could alter dramatically over those operating conditions, hence the concern for adaptive control. The latter may be loosely defined as *a system having a controller with adjustable parameters and a mechanism for adjusting the parameters* (Åström and Wittenman 1995). To “adapt” means to change oneself so that one’s behavior will conform to new or changed circumstances (Ioannou and Sun 1996). Such a capability mimics the learning capability in many sensorimotor systems. The internal model paradigm attempts to elucidate such a learning mechanism from the perspective of sensorimotor integration whereas adaptive control theory tackles an abstract learning problem in general based on mathematical analysis. From a behavioral perspective, the terms “learning” and “adaptation” are sometimes taken to connote, respectively, an active or voluntary construction of new actions as opposed to a passive or involuntary adjustment to a novel environment. Such semantic distinction is less relevant in an engineering context, however, where these terms are often used interchangeably.

The study of adaptive control over the past decades has evolved a strong mathematical foundation for designing and analyzing sophisticated learning systems (Slotine and Li 1991; Åström and Wittenman 1995; Ioannou and Sun 1996). On the other hand, the recent study of internal models has provided a rich empirical characterization of complex nonlinear physiological systems, which may not be always amenable to rigorous mathematical examination. It is beneficial if a clear connection could be drawn between internal model and adaptive control theory and between the varieties of internal models found in varying sensorimotor systems.

This paper provides a comprehensive engineering overview on adaptive control theory and biological overview on internal models in sensorimotor systems. Interweaving the two traditionally distinct disciplines together systematically across multiple sensorimotor systems leads to a unified conceptual framework that illuminates the applicability of adaptive control theory to the study of internal model, and *vice versa*. This interdisciplinary approach should facilitate the experimental elucidation of the mechanisms of internal model and the reverse engineering of such neural mechanisms into novel brain-inspired adaptive control paradigms in future.

2. Adaptive Control Theory of Internal Model

The internal model paradigm has its roots in the engineering theory of adaptive control and state estimation. This section gives a broad overview of some basic adaptive control and system identification techniques and concepts that are closely related to the physiological examples to be discussed in the later part of the paper.

2.1. Performance measures of adaptive control

Like classical control theory, the performance of adaptive control is generally assessed according to certain rigorous criteria such as stability, convergence and robustness. Stability means that when a system is sufficiently close to equilibrium, the system states can be kept arbitrarily close to the equilibrium point under perturbation and return to the equilibrium point when the perturbation is removed (Slotine and Li 1991). Convergence means that given a bounded input, the output will be bounded and tends to a steady state over time. Besides output convergence, convergence of the parameter estimates for the plant (i.e., the process to be controlled) and controller to their true values is also of major concern in adaptive control and system identification problems. Finally, an otherwise stable controller may become unstable in the presence of small disturbances or unmodeled dynamics. Robustness describes the amount

of such uncertainty the system can tolerate before controller performance is significantly compromised (Åström and Wittenman 1995; Ioannou and Sun 1996; Ogata 1997).

Such performance measures are fundamental to engineering adaptive control design but are not always included in the study of sensorimotor systems. Incorporation of such performance assessment in neural modeling is crucial in that it may aid in hypothesis testing, i.e., a proposed internal model paradigm may be rejected if it does not demonstrate the stability, convergence and robustness characteristics of the sensorimotor system under study.

2.2. Classification of adaptive control laws

An adaptive controller (whether biological or man-made) typically comprises three key elements: (1) a control objective J to be optimized adaptively; (2) a control law with adjustable parameters $C(t)$; and (3) an adaptation or estimation law $dC(t)/dt$ for adjusting $C(t)$ toward the optimal control objective (Åström and Wittenman 1995). All three elements are generally dependent on the input signal $r(t)$, control signal $u(t)$ and the output signal $y(t)$. Depending on the control system architecture and nature of the adaptation law, the adaptive controller may be classified as direct or indirect, and deterministic or stochastic.

Direct and indirect adaptive control—In *direct adaptive control*, the adaptation law is implemented directly whereas in *indirect adaptive control*, the adaptation law is implemented by first estimating some unknown plant parameters and/or state variables. Figure 1a shows the basic architecture of direct adaptive control. Here, the controller parameters $C(t)$ are adapted directly based on the input/output and control signals, with no explicit estimation of the plant parameters or state variables necessary. In contrast, in indirect adaptive control (Fig. 1b) the unknown plant parameters or state variables are continuously updated based on some parameter estimation or state estimation law. The resultant estimates are then used to compute the controller parameters at each time t . For indirect adaptive control, the design of the controller considers the plant parameter estimates or state variable estimates as the true parameters or state variables at each time t . This is called the *Certainty Equivalence Principle* and is a widely accepted assumption in adaptive controller design (Åström and Wittenman 1995; Ioannou and Sun 1996).

In general, indirect adaptive control is more flexible than direct adaptive control because the estimation law and control law can be designed separately (Ioannou and Sun 1996). However, stability and convergence are not guaranteed because the estimated plant model may not always satisfy controllability and stability conditions which are requisite for controller design (Ioannou and Sun 1996). Persistent excitation of the input signal is generally required for unbiased identification of the system parameters. However, identifiability could be lost under closed-loop conditions (Åström and Wittenman 1995). For direct adaptive control, parameterization of the plant model into the controller parameters is not necessary (Ioannou and Sun 1996). Also, stability and convergence can be guaranteed with proper design of the controller. However, designing an adaptive controller directly is a challenging task that is not always feasible.

The notion of direct or indirect adaptive control is implicit in sensorimotor modeling although its importance is often not well appreciated. In the context of behavioral neurobiology, indirect adaptive control may be indicated when cognition of movement dynamics or kinematics is involved such that the brain has a sense of these state variables before adapting the motor command. For example, in motor learning, neurons in the primary motor cortex (M1) and the supplementary motor area (SMA) (Li, Padoa-Schioppa et al. 2001; Padoa-Schioppa, Li et al. 2002; Padoa-Schioppa, Li et al. 2004) are found to encode the changes in movement dynamics. Such changes in neuronal activities represent the adaptation outcomes of kinematics-to-dynamics transformation, which could represent either direct or indirect adaptive control. On

the other hand, in vestibular motion sensing, neurons in the vestibular nuclei and the rostral fastigial nucleus of the cerebellum are found to compute a solution to the inertial motion detection problem faced by the otolith organs and semicircular canals (Angelaki, Shaikh et al. 2004). The identified kinematic variables may contribute to other motor control actions, for instance, posture control to maintain balance (Kuo 1995). The explicit system identification process prior to the implementation of motor control indicates an indirect adaptive control scheme.

Deterministic and stochastic adaptive control—A system is said to be *deterministic* when the effects of random disturbance and/or noise are negligible. For example, the certainty equivalence principle takes the estimated parameters as the true values for the purpose of adaptive controller design. In practice, sensorimotor systems are subject to random physiologic perturbations and such variability may influence internal model formation. *Stochastic adaptive control* takes into account the effect of such uncertainties during adaptation. This may be achieved directly or indirectly, i.e., through stochastic state and parameter estimation. For example, Kalman filter is a stochastic state estimator which attempts to minimize the covariance of the state estimate. Some stochastic adaptive controllers, such as Hebbian covariance feedback control, actually make use of such random disturbances to directly achieve adaptive control (see below).

2.3 State estimation and parameter estimation paradigms

System identification (i.e., estimation of unknown system parameters) and state estimation are important first steps in indirect adaptive control. Several deterministic or stochastic system identification schemes are pertinent to the modeling of sensorimotor systems.

Luenberger observer for state reconstruction in deterministic systems—The Luenberger observer reconstructs the state of a dynamic system from partial observations of the system's outputs (Luenberger 1965). A *full-order observer* reconstructs all the state variables whether they are directly observed or not (Fig. 2). A *reduced-order observer* reconstructs only a subset of the state variables. The observer usually assumes a similar form as the plant dynamics but with a correction term derived from a gain-adjusted difference between the actual plant output and the observer output. The feedback gain is chosen such that the error in the observer output converges to zero asymptotically.

In designing an observer, the criterion of *observability* is crucial. A system is said to be observable if the states at any time can be determined from the observation of the output over a finite time. An observer can be designed if and only if the observability condition is satisfied (Ogata 1997). While the Luenberger observer theory applies mainly to linear systems, recent work attempts to design observers for nonlinear systems. For instance, Zhao and Slotine have derived a discrete nonlinear observer for inertial navigation based on contraction theory (Zhao and Slotine 2005).

The theory of observer in dynamic systems is of relevance to the internal model paradigm in that the central neuronal encoding of dynamic and kinematic movement variables (Shidara, Kawano et al. 1993; Padoa-Schioppa, Li et al. 2002; Angelaki, Shaikh et al. 2004; Padoa-Schioppa, Li et al. 2004) is functionally analogous to an observer. The ability of an observer to reconstruct unobserved state variables over time may underlie some sensory systems' ability to resolve underdetermined sensory cues (Merfeld, Zupan et al. 1999; Angelaki, Shaikh et al. 2004). A deterministic observer may also provide a mechanistically much simpler alternative to stochastic state estimators such as Kalman filter, which has been proposed by some investigators as possible mechanism of forward (or predictive) internal model of sensorimotor integration (Wolpert, Ghahramani et al. 1995; Park and Kuo 2005).

Kalman filter for state estimation in stochastic systems—The Kalman filter for state estimation in stochastic dynamic systems (Brown and Hwang 1997) is a recursive algorithm based on the minimum mean-square error approach (Fig. 3). The state variable estimates are updated through Bayesian rule. For systems with linear dynamics and Gaussian-distributed measurement noise, the Kalman filter can be formulated nicely using efficient matrix operations on the mean and covariance estimates of the state variables. Kalman filter can be viewed as a stochastic counterpart of the Luenberger observer with an optimal choice of the gain matrix. For nonlinear systems, an extended Kalman filter which linearizes the system using first-order Taylor series expansions is often used. However, convergence of the extended Kalman filter is not always guaranteed and may require complex online parameter tuning (Wiberg, Powell et al. 2000).

A novel Bayesian estimator, called particle filter, overcomes the limitations of the Kalman filter in that it is generally applicable to *nonlinear* and *non-Gaussian* state-space models (Andrieu, Doucet et al. 2004). The particle filter approximates the posterior distribution of estimates by a large number of samples (particles), each with a normalized assigned weight. The filter updates the particle positions and the corresponding weights recursively at the cost of intensive computational effort. The study of particle filters is very new and is still evolving rapidly.

For all the recent excitement about the possible relevance of Kalman filter or other optimal filters as a sensorimotor state estimator (see section 5.2), there is as yet no direct evidence to indicate that central neurons are capable of the complex covariance computations that are basic to the Kalman filter update equations (Fig. 3). Furthermore, neurons are inherently noisy (but not necessarily Gaussian) in their operation. The variability in neurotransmission may actually overshadow the uncertainties in sensory measurements. If so, a Luenberger observer may be just as effective as a Kalman filter in learning the system state. In both cases, the observability criterion is required for any system state identification to be feasible. Nevertheless, it is possible that neurons could take advantage of the intrinsic variability of sensory information in order to up- or down-regulate synaptic transmission based on some covariance learning rule (see Fig. 12 below). Further studies are needed in order to elucidate the neuronal substrates underlying stochastic estimation and control in sensorimotor systems.

Parameter estimation—Parameter estimation is an important component in indirect adaptive control. The least-squares method is a basic technique for parameter estimation that does not require any prior information about the parameters or the measurement noise. For real-time applications, it is more convenient to use a recursive least-squares method which automatically updates the parameter estimates at each time step. To estimate a time-varying parameter, the estimation algorithm can be modified by incorporating a “forgetting” factor which weights the data in time (Åström and Wittenman 1995).

Maximum-likelihood is another technique commonly used for parameter estimation when the nature of measurement noise is known. The aim of maximum likelihood estimation is to find the parameter value(s) that makes the observed data most probable to occur.

The extended Kalman filter allows parameter estimation based on Bayesian statistics. The extended Kalman filter is itself a recursive algorithm and can be used for online parameter estimation (Brown and Hwang 1997).

2.4 Deterministic adaptive control paradigms

An adaptive control system can be designed in varying ways depending on the application. There are two facets to each paradigm: the control system architecture and adaptation law. The

following provides an overview of some deterministic adaptive control paradigms that may be applicable to the modeling of sensorimotor systems.

Self-tuning regulator—In classical feedback control, the controller design is determined as a function of the plant. When the plant model is unknown, the controller parameters are adapted using an (implicit or explicit) estimate of the plant model based on measurements of the instantaneous input, output and control signals. The estimation and adaptation procedures are performed online automatically. This construction is called “self-tuning regulator (STR)” to emphasize the automatic tuning of the controller. STR can be in direct or indirect style as shown in Figs. 1a and 1b. In STR, learning is unsupervised, i.e., without the need for an explicit “teacher” or reference model.

High-gain adaptive control—High-gain adaptive control (HGAC) is the simplest STR in the form of direct adaptive control. In classical feedback control of an unstable plant, a proportional controller with a sufficiently high gain can stabilize the closed-loop system even when the plant is nonlinear, provided the plant satisfies certain “strictly positive realness” (SPR) condition (Mareels 1984). This is in contrast to non-SPR systems, where high gain may lead to instability especially when the feedback has significant delays. This idea is employed in HGAC with a continuously adapting controller gain K (Fig. 4). Here, stabilizing the plant output $y(t)$ is equivalent to tracking a zero reference input, i.e., $r(t) = 0$. In this case the control and adaptation laws can be formulated, respectively, as follows:

$$u = -Ke \quad (1)$$

$$\dot{K} = \eta e^2 \quad (2)$$

where K is the controller gain, $e(t) = r(t) - y(t)$ is the error signal, and $\dot{K}$ is the adaptation rate. It can be easily verified that $e(t)$ will converge to zero asymptotically as K increases to some large but finite value. Estimation of the plant parameters is not necessary. Indeed, convergence is guaranteed even when the plant input-output relationship is nonlinear (but bounded). However, convergence is restricted to the reference input $r(t) = 0$ only. For any non-zero and/or time-varying reference input, $e(t)$ generally does not converge to zero. Moreover, this design scheme is not robust to measurement errors (Polderman and Mareels 2004).

The HGAC paradigm may underlie the remarkable ability of the motor system in stabilizing unstable dynamics, as reported by some authors (Burdet, Osu. et al. 2001) (see section 3.3).

Gain scheduling—Another simple direct adaptive control method is to adjust the controller parameters in a style of lookup table called *gain scheduling* (Fig. 5). The lookup table serves as a reference model for supervised learning. Consider a system with N distinct operation points. The operation points are measurable variables that correlate with changes in the system dynamics. For instance, in an aircraft system the operation point can be defined by a certain range of, say, Mach number and altitudes over which the flight system dynamics remains constant. For each of these operation points, a separate controller is designed to meet the performance requirement. A switching mechanism picks the controller specific to each operation point encountered during operation. The switching mechanism can be discrete or a continuous function of the operation states. Since the controller parameters are computed offline, this method is capable of handling fast variations in plant parameters and is only limited by the speed of detecting the operation points. However, operating cost increases as more operation points are included. Moreover, frequent switching between different controllers could lead to instability. Such a design is difficult to analyze and stability is usually verified through simulation. Another drawback is that the calculation of controller parameters is done

entirely offline *a priori*. The lack of online adaptation can lead to poor performance or even failure in the case of unexpected changes. Despite its disadvantages, the simplicity of gain scheduling makes it popular in some applications such as flight and process control (Åström and Wittenman 1995).

Gain scheduling may underlie the passive (non-adaptive) component of motor control (Flash 1987) (see section 3.1) or some aspects of posture control (Kuo 1995).

Model reference adaptive control—In model reference adaptive control (MRAC) a reference model is given that characterizes the desired properties of the system, which is used to supervise the adaptation of the controller. The objective of MRAC is to adapt the closed-loop system toward the reference model, such that the difference between the plant output, $y(t)$, and the reference model output, $y_m(t)$, converges to zero over time. MRAC can be implemented in the style of either direct or indirect adaptive control. Figure 6 shows a schematic of a direct MRAC. The controller continuously adapts in response to the error signal, $e(t)$. For linear plant dynamics, stability and convergence of error are usually guaranteed with a proper choice of the control and adaptation laws. A commonly used MRAC adaptation law is the *delta rule*,

$$\frac{dC_1}{dt} = -\gamma_1 e(t) r(t) \quad (3)$$

$$\frac{dC_2}{dt} = -\gamma_2 e(t) y(t) \quad (4)$$

where C_1 and C_2 are the controller gains to be adjusted, $r(t)$ and $y(t)$ are the inputs to the feedforward and feedback controllers respectively, γ_1 and γ_2 are the adaptation rates, and $e(t)$ is the output error between the reference model and the closed-loop system (Åström and Wittenman 1995). Stability of the delta rule is guaranteed for linear plant dynamics according to hyperstability theory (Slotine and Li 1991; Åström and Wittenman 1995). This delta rule for MRAC adaptation has been applied to other supervised learning problems such as feedforward neural networks (McClelland and Rumelhart 1988).

Although MRAC is primarily an adaptive control design, it can be readily reconfigured as an observer to learn a forward model. In MRAC, the closed-loop system (those inside the dotted box in Fig. 6) converges to the reference model. To configure it as an observer, the closed-loop system can be made to converge to the unknown plant, i.e., the roles of the “reference model” and the “unknown plant” are reversed. By the principle of duality, observability of the reference model (the new “unknown plant”) is contingent on its controllability. A system is said to be controllable if any arbitrary states can be achieved by applying an admissible input (Ogata 1997). Controllability is typically applicable only to linear systems. Nevertheless, *differential geometry* provides the mathematical tools to define a similar condition for nonlinear systems (Slotine and Li 1991). Furthermore, this condition determines whether a nonlinear system is feedback linearizable. *Feedback linearization* is an approach to transform a nonlinear system into an equivalent linear system. Such geometric control approach is useful in designing robust and adaptive nonlinear controllers (Slotine and Li 1991; Nelson 1999; Anzaldo-Meneses 2002). Although the controllability condition guarantees convergence of output error, it should be noted that convergence in output error does not guarantee convergence of the model to the unknown plant. Generally, convergence of system parameters requires sufficiently persistent input excitation.

Adaptive sliding control—Slotine and Li (Slotine and Li 1987; Slotine and Li 1987) have presented a STR algorithm (Fig. 7) which consists of a proportional-differential feedback loop and a full dynamics feedforward loop, with the unknown system parameters estimated online. By assuming a linearly parameterized dynamic model, the adaptive controller is proven with Lyapunov stability analysis to be globally stable with zero tracking errors. This is achieved by introducing a time-varying sliding surface, s , defined as,

$$\dot{s} = \dot{q}_r = \dot{q} - \dot{q}_r = \dot{\tilde{q}} + \Lambda \tilde{q} \quad (5)$$

$$\tilde{q} = q - q_d \quad (6)$$

where $\tilde{q}$, $\dot{\tilde{q}}$ are the error signals in position and velocity, q_r is a “reference” trajectory acting as an intermediate variable to compute the sliding variable s , and Λ is a symmetric positive definite matrix. The control law is defined as,

$$u = \hat{Y}\hat{a} - K_D s \quad (7)$$

where u is the control signal, $Y = Y(q, \dot{q}, \ddot{q}, \ddot{q}_r)$ is generally a matrix function, $\hat{a}$ is a vector of the parameter estimates, $\hat{Y}\hat{a}$ denotes the estimated inverse dynamics of the plant and serves as the feedforward controller, and K_D is the gain of the feedback controller.

Lyapunov stability analysis then guarantees that the errors ($\tilde{q}$ and $\dot{\tilde{q}}$) converge asymptotically to the sliding surface $s = 0$, which corresponds to zero tracking errors in both velocity and position.

Although adaptive sliding control is unsupervised, the introduction of a reference trajectory is in some respects similar to MRAC. The sliding variable describes the dynamics of a reference model. Adaptation is achieved by adjusting the parameters in the feedforward control loop, which identifies a coarse model for the inverse dynamics of the plant. The algorithm provides a computationally simple method for designing an STR controller that is stable and robust with fast convergence even for nonlinear plant dynamics. The need for explicit estimation of the inverse of plant dynamics means that it is an indirect adaptive control paradigm. However, the inverse dynamics itself directly provides the control law and so it may also be considered a direct adaptive control method. This algorithm has been successfully implemented on a robot manipulator with a large load of unknown properties (Slotine and Li 1986b; Slotine and Li 1987). The control effort is shown to be promising with precise and rapid parameter estimation capability. Recently, the algorithm has been extended to tackle control problems with uncertain kinematics and the resultant algorithm is capable of adapting both unknown kinematics and dynamics parameters online (Cheah, Liu et al. 2004; Cheah, Liu et al. 2005).

Adaptive sliding control has important implications in providing a rigorous engineering basis for the feedback error learning paradigm (Fig. 8) proposed by Kawato et al. (Kawato, Furukawa et al. 1987) for inverse internal model adaptation during motor learning (see section 3.2).

Multi-model adaptive control—Multi-model adaptive control (Fig. 9) is MRAC and gain scheduling combined. A set of reference models corresponding to differing operation points is given *a priori*. A controller is designed for each of the reference models. At every instant, the outputs of the reference models are compared with the actual output of the plant. An optimal reference model that best fits the plant is then selected by a switching mechanism to compute

the control input. The control performance may be further improved by introducing an adaptive mechanism in the reference models. Like gain scheduling, convergence and stability of multi-model adaptive control are difficult to analyze and the cost of design grows rapidly with number of operation points.

Multi-model adaptive control may underlie the multiple paired forward and inverse models for motor control proposed by Wolpert and Kawato (Wolpert and Kawato 1998) (see section 3.4).

2.5 Stochastic adaptive control paradigms

Generalized predictive control/Model predictive control—Predictive control is a form of control which incorporates a prediction of the future behavior of the system into its formulation. Predictive control is most useful for systems whose future behavior could depart substantially from that perceived at the present time, for instance, time-delayed systems, poorly damped and non-minimum phase systems, etc. In particular, model predictive control (MPC) (also known as the receding-horizon control (RHC)) makes explicit use of a model of the plant in deriving a control signal that optimizes the objective function over a finite number of future time steps (the planning horizon). Only a portion of the calculated control input is implemented (the execution horizon). Then, the process repeats.

Generalized predictive control (GPC) is a general form of model predictive control and is simple, effective and widely used in many industrial applications (Huang, Tan et al. 2002). GPC implements a single-step execution horizon and uses a quadratic objective function which is characterized by the sum-of-squared difference between the estimated states variables and a desired reference, and the cost of the control efforts. By incorporating an appropriate disturbance model, one may take into account behaviors which are not reflected by the process model, for instance, non-measurable input, noise and model errors, thereby increasing the robustness. The theoretical basis of GPC has been widely studied. Analytical solutions are available (in the absence of constraints) (Camacho and Bordons 2004). Inclusion of constraints is simple and stability can be guaranteed with suitable modifications to the basic algorithm (Huang, Tan et al. 2002).

The predictive control paradigm may be useful in elucidating some aspects of motor control, such as motor planning (Sober and Sabes 2003). The RHC has been successfully implemented in path planning of complex multiple autonomous mobile robot system (Kuwata and How 2004). Motor planning is essentially a path planning problem as well. The successful implementation of RHC in robotic systems suggests the possibility of describing motor planning in physiological systems using a similar approach.

Reinforcement learning—Reinforcement learning (Sutton and Barton 1998) is a “neurally-inspired” artificial intelligence paradigm which, at its core, is analogous to stochastic adaptive control. Here, an agent tries to maximize certain reward function as a result of its actions. As with most forms of machine learning, the agent is not told which actions to take but instead must discover through trial-and-error which actions yield the maximal expected reward (which could come instantaneously or in future). On a stochastic task, each action must be tried many times in order to gain a reliable estimate of its expected reward.

The four key elements of reinforcement learning are the *policy*, *reward function*, *value function*, and, optionally, *model* of the environment (Fig. 11). Each element has a corresponding counterpart in adaptive control. The *policy* at any system state specifies the probabilities of various possible actions at time t . These probabilities are updated after each trial based on the resulting rewards. The policy in reinforcement learning is analogous to the *control law* in stochastic adaptive control, since it relates the control input (the action) to the states of the system, and can be adapted during the operation. A *reward function* quantifies each policy's

expected reward or penalty as a function of time (present and delayed). In terms of adaptive control, the reward can be regarded as the instantaneous values of the objective function to be optimized. The *value function* is the accumulation of the expected reward over time resulting from a given policy, starting from the present state. Hence, it is an estimation of the “value” of that particular state. This coincides with the cumulative objective function in adaptive control. Unlike other forms of supervised learning, reinforcement learning does not require explicit information about the gradient of the value function.

The *model* mimics the behavior of the environment. Given a state and action, the model might predict the resultant next state and next reward. Early reinforcement learning methods used solely trial-and-error rather than an estimated model of the environment to update the policies. However, it gradually became clear that reinforcement learning methods are closely related to dynamic programming (Bertsekas and Tsitsiklis 1996), which does use models, and that an intermediate model estimate may contribute to the reinforcement learning process. In this respect, traditional and modern reinforcement learning methods are analogous to direct and indirect adaptive control.

Learning from reward and punishment is an essential element of human and animal behaviors. A reward (e.g. food) can reinforce certain motor response by classical conditioning. If a conditioned stimulus is paired with this reward over time, a motor response can be elicited by the conditioned stimulus alone. The neurotransmitter *dopamine* is known to be involved in reinforcement of human and animal behaviors (Schultz, Dayan et al. 1997). The response of the dopamine neurons replicates the function of *temporal-difference* (TD) error in the context of reinforcement learning (Schultz, Dayan et al. 1997; Doya 1999; Doya, Kimura et al. 2001). Simulation results show that TD error model is capable of reproducing the prediction of reward given the conditioned stimulus after training (Schultz, Dayan et al. 1997; Suri 2001). The TD model has been modified by Redish to elucidate the neurophysiological mechanism of drug addiction (Redish 2004). By incorporating a noncompensable drug-induced dopamine increase to the normal TD function, the model provides a computational basis for the learning mechanism that over-selects actions leading to drug receipt.

Hebbian Feedback Covariance Control—Hebbian synaptic plasticity was first postulated by Hebb over 50 years ago as a mechanism of learning and memory (Hebb 1949). The classical Hebbian model (Hebb 1949) states that the strength (or gain) C of a synaptic connection is modified according to an adaptation law of the form:

$$\frac{dC}{dt} = k(y \cdot u) \quad (8)$$

where y , u are mean firing rates of input and output neurons, respectively, and k is an adaptation constant. This synaptic adaptation law has been widely taken as the basis of NMDA receptor-dependent long-term potentiation (LTP) in some hippocampal and neocortical neurons.

However, there are some limitations to the classical Hebbian model, not the least the inevitable occurrence of runaway instability and irreversible saturation resulting from sustained or random coactivity of interacting neurons, leading to difficulties in its implementation (Poon 1996; Young 1997; Young and Poon 2001). Moreover, the dependence of the classical Hebbian model on pre- and post-synaptic activities local to the adapting neuron does not lend itself to feedback control applications. To circumvent these difficulties, a stochastic synaptic adaptation law called *Hebbian feedback covariance adaptation* has been proposed (Young and Poon 2001). Instead of pairing the mean input and output activities of the controller neuron, the new adaptation law modifies the synaptic strength by correlating the temporal variations of the pre- and post-synaptic neural activities about their respective mean values. Here, the pre- and post-

synaptic activities of the controller neuron correspond to the feedback signal and control signal, respectively. This adaptive control paradigm can be viewed as a reinforcement learning system driven by spontaneous, random perturbations in the control and feedback variables (Fig. 12). The new adaptation law becomes,

$$\frac{dC}{dt} = k_1(\delta y \cdot \delta u) - k_2 C g(\delta y, \delta u; y, u) \quad (9)$$

where $g(\cdot)$ is some positive definite function, which acts as a decay term to avoid saturation. Depending on the sign of the constant k_1 , the algorithm can describe synaptic LTP (for $k_1 > 0$) or LTD (long-term depression) (for $k_1 < 0$).

The Hebbian feedback covariance controller is stochastic in nature and employs a direct adaptive control approach. The controller gain, C , is updated according to the input/output relationship of the system, without explicit estimation of the plant parameters. In addition, the adaptive controller can be designed in such a way to optimize a certain objective function (which is essentially the Lyapunov function (Slotine and Li 1991)). By applying the Barbalat's lemma (Slotine and Li 1991), the objective can be guaranteed to converge to a minimum value at steady state.

The discussions of the algorithm so far have focused on static input-output relationship only. In reality, systems are usually governed by dynamical relationships as a result of, for instance, slow time constants and time delays. Young and Poon (Young and Poon 2001) modified the original algorithm for dynamical systems by introducing a near-term objective function, Q . The Hebbian covariance feedback law is applied to the near-term objective function which, with some suitable transformations of the state variables, will lead to minimization in the long term.

To illustrate, consider a first-order nonlinear system,

$$\dot{y} = q(y, u) \quad (10)$$

The steady state solution is obtained as,

$$y_0 = f(u_0) \quad (11)$$

An intermediate variable, z is defined as a filtered version of the original state variable y ,

$$z(y, \dot{y}, u) = f(u) \quad (12)$$

such that $z \rightarrow y_0$ as $f(u) \rightarrow f(u_0)$. Hence, the system becomes static in z , and the same algorithm can be applied on this intermediate (or filtered) state variable with a near-term objective function defined in z and u . In steady state, the near-term objective function converges to the long-term objective function. The same approach is generally applicable to systems of higher order.

The Hebbian feedback covariance control is computationally simple compared to conventional self-tuning or model-reference adaptive control, which generally require a prescribed reference model or desired trajectory. On the other hand, it is also a form of reinforcement learning with an implicit reinforcement signal, namely the covariance of the filtered state and the control signals. Furthermore, optimality is guaranteed through Lyapunov analysis, whereas GPC is generally suboptimal.

The Hebbian covariance self-tuning optimal control was first inspired by the modeling of the adaptive optimal control in the respiratory system (Fig. 13) (Poon 1996; Young and Poon 1998) (see section 4.2).

3. Deterministic Adaptive Control in Motor Learning

Humans demonstrate a remarkable ability to generate accurate and robust motor behavior under varying conditions, even in the presence of uncertainty. This capability is often not matched in robotic control systems. The following reviews the classical equilibrium-point hypothesis and the current internal model hypothesis of motor control from the perspectives of adaptive control theory commonly used in robotics and other engineering applications. Possible associations between these biological and engineering paradigms are examined and discussed. Comparison of the equivalent neural and engineering paradigms of adaptive control will bring to light the general design principles underlying both biological and man-made learning systems.

3.1. Equilibrium-point hypothesis and impedance control as gain-scheduling

Early studies suggested that certain motor tasks such as hand reaching movement may not necessarily require adaptation. According to the equilibrium-point hypothesis, the motor command may modulate the stiffness and viscosity of the muscles which act as a spring to “pull” the hand back to some equilibrium positions of the muscles regardless of external disturbances. Impedance control is an extension of the equilibrium-point hypothesis in the context of engineering robotics, where robotic control is achieved by regulating its output impedance which characterizes the dynamic behavior of the robot at the port of interaction with the environment (Hogan 1985; Hogan and Buerger 2005). Output impedance is intrinsic to the robotic control system and does not require precise knowledge of the environment (Hogan and Buerger 2005). This gives impedance control its appeal since in general a complete model of the environment is lacking and the environment may vary from task to task. Similar impedance control principles have been applied to the “negative impedance” design of a synergistic patient-machine interface for assisted mechanical ventilation (Poon and Ward 1986; Poon, Lebowitz et al. 1997) and the “dynamic clamp” implementation of artificial ion channel dynamics in neuron-computer interface (Prinz, Abbott et al. 2004).

Because the neuromuscular system is spring-like by nature, where the impedance can be modulated by coactivation of antagonist muscles without changing the resultant output torque (Hogan 1984), the intrinsic impedance of the muscle system forms the basis of the equilibrium-point hypothesis. Impedance control theory helps to explain the control mechanisms implemented by the central nervous system in, for instance, maintaining stability in various force-production tasks (Rancourt and Hogan 2001). Experience from robotic control indicates that, due to the high nonlinearity and high degrees of freedom of the system, a complex computation effort is required for the inverse dynamics calculation in order to generate the necessary motor commands to realize a desired trajectory of the limb faithfully. The equilibrium-point hypothesis suggests that such a high dimensional problem can be transformed into a set of low-dimensional equilibrium points (in the hand coordinates) where the elastic forces generated by all the muscles sum to zero (Abend, Bizzi et al. 1982; Bizzi, Accornero et al. 1984; Flash 1987; Bizzi and Mussa-Ivaldi 1990). Hence, the hand can be driven back to the equilibrium point through the action of the elastic force whenever it deviates from this point. By shifting this equilibrium point in space, the hand can be moved from target to target. It has been shown that hand movements in some reaching experiments can be closely described by the equilibrium trajectory (Flash 1987; McIntyre and Bizzi 1993).

Because stiffness of the arm under static condition is highly dependent on joint positions (Mussa-Ivaldi, Hogan et al. 1985), it may be inferred that the stiffness is *pre-calculated* at each

hand position. As such, the equilibrium-point hypothesis resembles the gain scheduling scheme of adaptive control (Fig. 5) with an implicit impedance control law given by the stiffness and viscosity of the arm muscles during movement and a continuous gain-switching schedule that is determined by hand position. This form of impedance control is analogous to a PD controller with a resultant torque given by the deviations of the instantaneous hand position and velocity from their corresponding virtual trajectories:

$$\tau = K(\theta_o - \theta) + B(\dot{\theta}_o - \dot{\theta}) \quad (13)$$

where K is the stiffness matrix, B is the viscosity matrix, and θ_o is the virtual trajectory for impedance control, which is equivalent to the reference trajectory for PD control (Hogan and Buerger 2005).

More recent studies argue against the equilibrium-point hypothesis, especially during high-speed movement (Flash 1987). Gomi and Kawato (Gomi and Kawato 1996) suggested that hand trajectory in a similar reaching experiment under low joint stiffness around the resting value did not follow the same equilibrium trajectory profile: a more complex trajectory was observed. However, the result is controversial since they used an oversimplified model for the neuromuscular system (Gribble, Ostry et al. 1998; Hodgson and Hogan 2000). Instead, Hodgson and Hogan (Hodgson and Hogan 2000) showed that a model-independent attractor trajectory could be verified experimentally, in agreement with the equilibrium-point hypothesis. In another set of reaching experiments, subjects were asked to move their hand between target points under the influence of an external force field that distorted the movement (Shadmehr and Mussa-Ivaldi 1994; Franklin, Osu et al. 2003). After a number of trials the subjects were able to modify their hand trajectory to converge to the original straight line path. After the learning, the external force field was removed and the subjects were asked to perform the same task. The result showed a “mirror image” of the hand trajectories compared to those during the learning process. This aftereffect provides evidence for the existence of an internal model which persists in the brain after the adaptation.

3.2. Feedback error learning as adaptive sliding control

Neilson et al. (Neilson, O'Dwyer et al. 1988; Neilson, Neilson et al. 1992) suggested that motor learning may be modeled as an internal model representation by some adaptive filter networks in the brain. Kawato et al. (Kawato, Furukawa et al. 1987) proposed a *feedback error learning* model as a possible mechanism of adapting internal models. The model comprises a feedforward and feedback loop (Fig. 8). The output of the feedback controller is used as an error signal to train the feedforward controller, which, at the completion of learning, will identify the inverse dynamics of the plant model. Thus, feedback error learning produces an inverse internal model. The error signal is kept in check by the error feedback loop during learning. The feedback error learning mechanism employs a direct self-tuning control scheme without explicitly estimating the parameters of the plant. The feedforward controller is tuned using the gradient method (Åström and Wittenman 1995) in order to minimize the error (Miyamura and Kimura 2000; Doya, Kimura et al. 2001). Computer simulation and experiments with PUMA robot arm have shown that this mechanism can successfully learn an internal model that gradually takes the place of the feedback controller. In support of this theory, Imamizu et al. (Imamizu, Miyauchi et al. 2000) found that an area near the posterior superior fissure contains neurons that maintain a high activity after visuomotor learning, and that a wide region of the cerebellum encodes an error signal that may guide the learning of the internal model. These regions may constitute the feedforward controller in the feedback error learning model.

It is of interest to note that the feedback error learning scheme actually shares a similar architecture as adaptive sliding control first proposed by Slotine et al. in the engineering literature (Slotine and Coetsee 1986a; Slotine and Li 1987; Slotine and Li 1987; Slotine and Li 1991). Adaptive sliding control implements a control law which combines a fixed feedback controller (PD controller) and an adaptive feedforward controller (an inverse dynamics of the system), which has a similar structure as feedback error learning. Nevertheless, the two algorithms use different adaptation laws: adaptive sliding control uses the sliding variable, whereas feedback error learning uses a neural network controller. Hanneton et al. (Hanneton, Berthoz et al. 1997) proposed that the experimental data of human control of movement are consistent with the predictions of adaptive sliding control.

Although the original formulation of feedback error learning relied largely on computer simulations (Kawato, Furukawa et al. 1987), mathematical proof for the conditions of convergence was subsequently considered based on Lyapunov analysis assuming an artificial neural network as the feedforward controller and “strongly stationary stochastic processes” as the inputs (Gomi and Kawato 1993). Convergence is contingent on the stringent condition that the learning rate is very small and positive definite. On the other hand, adaptive sliding control has a rigorous mathematical foundation based on Lyapunov analysis such that stability, convergence and robustness are generally guaranteed for well-defined nonlinear plant dynamics. Moreover, the speed of learning can be arbitrarily fast and is limited only by extraneous factors such as unmodeled high-frequency dynamics, time delay and sampling time. The similarities and differences between the adaptive sliding control structure (Fig. 7) and feedback error learning structure (Fig. 8) underscore the importance of applying multidisciplinary neuroscience and engineering approaches to the solution of such complex problems.

3.3. Dual sliding mode and high-gain adaptive control of hand movement

Although the internal model and impedance control hypotheses were initially thought to be contradictory to one another (but see (Mussa-Ivaldi and Bizzi 2000)), recent studies suggest that they could co-exist during motor learning (Burdet, Osu. et al. 2001; Franklin, Osu et al. 2003). Electromyographic (EMG) recordings of the arm muscles during the reaching experiment show that muscle activations increase rapidly in the early phase of the learning as a result of agonist-antagonist co-contraction. After a couple of trials, the activations start to decay exponentially, but with a slower time constant (Franklin, Osu et al. 2003). This suggests that some adaptation exists in adjusting the impedance of the muscles in addition to shaping the internal model. In the early phase of the learning, the rapidly increased impedance reduces the effect of the destabilizing external disturbance to provide a better tracking of the desired trajectory. The hand-path error is reduced at about the same rate as the increase in the EMG signal, suggesting a direct correlation of the impedance control with the adaptation of hand path. Thus, whereas an internal model may eventually learn the dynamics of the arm and the environment, an up-regulated impedance controller in the early phase of learning may provide necessary stability that facilitates the learning of the internal model. Concurrent with this learning process, there appears to be an optimization process which attempts to minimize metabolic activity by reducing the activation of muscles. As the learning of the internal model proceeds, the need for the impedance controller becomes less critical and hence, the activations of the muscles are gradually relaxed in order to minimize metabolic cost.

The combined adaptation of impedance control and inverse-dynamics internal model learning suggests a dual adaptive control scheme across experimental trials. The joint stiffness provides the gain between the joint torque and the deviation from the equilibrium point. This gain grows rapidly in the first few experimental trials when the feedback error is large. The increase in gain over movement trials is in agreement with high-gain adaptive control, which is known to

provide stability (Fig. 4). Thus, whereas impedance control of arm movement may be likened to continuous gain-scheduling within a movement trial, the increase in arm impedance across movement trials in learning an unstable task is analogous to high-gain adaptive control.

In subsequent trials, the internal model gradually adapts and dominates over this high-gain adaptive controller. As before, the adaptation of the inverse internal model is similar to adaptive sliding control. Such a dual controller is expected to outperform either of these controllers since the latter are able to complement each other in order to achieve a better control effort.

3.4. Modular internal models as multi-model adaptive control

Humans are capable of learning a large number of skills, each of which may correspond to a different internal model. However, many of these skills have some common characteristics (compare, for instance, playing tennis and table tennis) and it is likely that the human brain has a more effective way to organize the internal models than simply store one model for every single task.

The idea of modular control architecture has drawn much attention in the study of motor control in human and humanoid robots (Kawato, Furukawa et al. 1987; Kawato 1999; Doya, Kimura et al. 2001; Schaal 2003). It is proposed that instead of a single gigantic model with a large number of adjustable parameters, multiple smaller scaled models co-exist in the brain with each being responsible for only a small subset of actions. Several lines of evidence support such an architecture (but see (Berniker 2000)). In the motor learning experiment, the de-adaptation to the normal condition and re-adaptation to a previously learnt condition are much faster than adaptation to a new condition. This suggests that humans retain multiple “learnt” internal models and are able to switch among them on the fly (Doya, Kimura et al. 2001). Secondly, Mussa-Ivaldi et al. (Mussa-Ivaldi, Giszter et al. 1994) have shown that simultaneous stimulation of two distinct sites in the frog’s spinal cord leads to vector summation of the forces at the ankle as generated by stimulating each site separately. This indicates that the total output is a linear combination of multiple separate signals, perhaps from multiple actuations in the system. Thirdly, functional magnetic resonance imaging (fMRI) studies on motor control have provided evidence of a multiple model architecture in the cerebellum (Imamizu, Kuroda et al. 2003). Activation related to learning of a specific task was observed in multiple regions in the brain, which suggests that there are probably multiple models working in parallel to learn the new task. Moreover, small common regions are activated by learning of different tasks. This may suggest that internal models acquired in these regions contribute common characteristics of these tasks. Finally, Osu et al. (Osu, Hirai et al. 2004) have shown that humans can adapt to two conflicting environments simultaneously. This suggests that the adaptation may involve simultaneous acquisition of multiple internal models and a mechanism for switching between models.

The mechanism of modular control architecture is akin to *gain scheduling* (and also *multi-model adaptive control*) (Figs. 5 & 9). Instead of a “hard” switching in gain scheduling which picks only one single controller at a time, the modular control scheme employs a “soft” switching mechanism in which each forward and inverse model pair is associated with a weight and the output is the weighted sum of the outputs of all these models (Wolpert and Kawato 1998; Kawato 1999; Doya, Kimura et al. 2001). Furthermore, unlike gain scheduling where the controllers are designed offline, learning for estimation and control are incorporated online in modular control. Wolpert and Kawato proposed a modular control architecture called *multiple paired forward and inverse models* (Wolpert and Kawato 1998). Each of the modules comprises a pair of forward and inverse models of the system. The output of each of these pairs is scaled by a weighting factor called *responsibility signal* which is updated by the difference between the actual and estimated state of the system. The gradient method and the feedback error learning are suggested as the adaptation laws for updating the state estimation and learning

of the controller respectively. This also means that an indirect approach with explicit estimations of the system states is probably employed in the learning.

The modular control scheme provides a novel architecture for describing human motor control. Such a scheme has been applied to humanoid robots (Schaal 2003; Schaal, Peters et al. 2004). The movement of the robot is divided into different dynamic movement primitives (DMP), which are basic units of action in the form of stable nonlinear attractors. It has been demonstrated that bipedal robots equipped with this modular control scheme can learn walking stably from the walking data of human subjects (Nakanishi, J et al. 2003).

Despite the flexibility of the modular control scheme, rigorous proofs for stability, convergence and robustness are lacking and performance is usually analyzed with computer simulations like in gain scheduling (but see (Lohmiller and Slotine 1998; Slotine and Lohmiller 2001; Slotine 2003)). Also, it remains unclear whether explicit estimation of state or a direct adaptive control mechanism is actually used in such an internal model. Furthermore, the question of how the brain defines the primitives of movement as with robotic control is unclear, since the architecture does not lend itself readily for inserting new model primitives into the hierarchy. Further studies will provide more insights for understanding the relevance of modular control to motor learning in humans.

4. Stochastic Adaptive Control in Motor Learning

Sensory measurements are inherently noisy, both in engineering and physiological systems. Recent evidence indicates that motor learning is not purely deterministic but may involve a probabilistic component. Two current paradigms of stochastic motor control based on Bayesian statistical inference and Hebbian covariance learning are reviewed.

4.1. Bayesian learning in sensorimotor systems

Humans are capable of adapting their behavior to allow for uncertainties. Learning an internal model alone may not guarantee optimal performance in the face of such uncertainties. For instance, in the classic dart-throwing experiment (Thach, Goodkin et al. 1992), there are infinite combinations of velocity and angle of throw that could hit the target. However, everyday experience shows that people tend to constrain the combinations within a small subset of possibilities. Even under microgravity where the only allowable trajectory is a straight line, there are still numerous possible velocities of throw to hit the target (Tryfonidis 1999; Poon, Tryfonidis et al. 2000). However, subjects might tend to throw at a low velocity to improve precision, as characterized by Fitts' law (Fitts 1954). Thus, the goal of sensorimotor adaptation may not be merely to hit the target (which is an ill-posed problem by itself), but to do so with great certainty.

In a target reaching experiment conducted by Kording and Wolpert (Kording and Wolpert 2004), human subjects were asked to reach a visual target in a virtual-reality setup with a laterally displaced visual feedback of their index finger. The subjects first learnt a prior probability distribution of the displacement. The visual feedback of the finger did not come in until midway through the movement. The reliability of the feedback was manipulated by applying varying levels of blurring. Results show that the relationship between the lateral shift and the deviations from target depends explicitly on the prior distribution and the uncertainties of feedback, consistent with predictions from Bayesian statistics. It appears that the brain is capable of deriving an optimal estimate of movement by combining the prior probability distribution of the estimate and instantaneous information from sensory feedback. This finding suggests that a probabilistic internal model may develop during sensorimotor learning. Thus, the objective of task performance is not only to minimize the mean errors, but also to minimize the variance of the errors. Similar Bayesian statistical inference has also proved useful in

explaining results of perception in a number of human psychophysical experiments (Knill and Richards 1996; Rao, Olshausen et al. 2002).

This Bayesian estimation scheme of Kording and Wolpert based on Gaussian assumption (Kording and Wolpert 2004) is essentially the same as a Kalman filter but in a static setting. In the context of Kalman filter, prior information is captured by the matrix P_0 , which computes the covariance of estimation. The estimation would converge generally for any positive definite matrix P_0 not equal to zero. Hence, even a simple initial guess would lead to a converging estimation although the goodness of this initial guess may affect the rate of convergence.

4.2. Hebbian feedback covariance learning model of respiratory motor control

The respiratory control system is a specialized motor system. The system is comprised of mechanical and chemical plants controlled by respiratory neurons in the medulla and pons regions of the brainstem. The recent discovery of various synaptic plasticity as well as neural adaptation and memory mechanisms in the central respiratory controller (Poon and Siniaia 2000) directly challenges the classical respiratory control models that are based on reflexogenic assumptions. In addition, many physiological studies support the optimal nature of breathing movements in ways that classical reflexogenic models fail to predict (Poon 2000). Such “intelligent” behaviors suggest that the respiratory controller could solve complex optimization problems and adapt to novel conditions. Inspired by such an intelligent control system, a Hebbian feedback covariance adaptive control paradigm which conforms to the neurophysiological system (Fig. 13) was proposed based on the Hebbian covariance learning rule (Fig. 12). The synaptic weight that determines the optimal input-output relationship is computed by correlating the corresponding intrinsic fluctuations that are ubiquitous in physiological signals (Young and Poon 1998). The basic Hebbian feedback covariance learning algorithm can be extended to account for system dynamics in the feedback loop (see section 2.5). Hebbian feedback covariance control has been successfully applied to the modeling of the respiratory system to predict the optimal adaptation behaviors during exercise and CO₂ inhalation (Young 1997; Young and Poon 1998). Moreover, robustness to noise disturbances is also verified with simulations. Further experimental evidence is required to identify the neural correlates of such an algorithm in respiratory control.

Hebbian feedback covariance control shows an example in which reverse engineering is applied to a physiological system toward developing a biologically-inspired engineering control paradigm. Such a paradigm may transcend its biological counterpart to suggest a general engineering control theory that is applicable to any practical system.

5. State Estimation in Sensorimotor Systems

System identification and/or state estimation are the first steps in indirect adaptive control. In some sensorimotor systems, sensory state estimation is a discrete cognitive process that precedes the resultant motor response. In vestibular sensory integration, for example, distinct perception of the kinematics of motion determines postural balance. In other sensorimotor systems, such as hand movement, the distinction between sensory integration and motor response may not be as clear-cut as the integration of sensory inputs could occur partly in the motor control system rather than the sensory or cognitive system alone. Such ambiguity accounts for the difficulty in distinguishing between forward models (for state estimation) and inverse models (for adaptive control) from psychophysics experiments as noted by some investigators (Karniel 2002).

5.1. Forward internal model for sensorimotor prediction

In motor control, the inherent delay of feedback control significantly deteriorates motor performance. Hence, predicting the future state of the motor system is essential for learning new tasks. The prediction is believed to be achieved through constructing a forward internal model, which captures the direct dynamics of the system. Detailed reviews of experimental evidence in support of forward model in motor control can be found in (Wolpert and Miall 1996). Sensorimotor prediction may generalize across different motor systems as well. Vercher et al. showed an eye-hand co-ordination for self-moved target tracking in which the eyes predict the movement of the hands with a forward model (Vercher, Sares et al. 2003). The process of motor prediction is equivalent to state estimation as discussed in section 2.3. Hence, the forward model could be formulated as a Luenberger observer or MRAC observer (both deterministic) or Kalman filter (stochastic) (Wolpert and Miall 1996). It should be noted that the forward model, as a state estimation model, may appear as part of an indirect adaptive control scheme (see section 2.2) that is separate from the inverse model.

5.2. Internal model for resolving vestibular sensory ambiguity

A critical step in human self-motion perception and spatial awareness is the integration of information from multiple sensory organs. One example can be found in identifying the actual motion associated with linear acceleration sensed by the otolith organs in the inner ears. Einstein's equivalence principle states that all linear accelerometers measure the total gravito-inertial force (GIF) which is the difference between gravitation force and inertial force due to linear acceleration. Same GIF measurement may result from different linear acceleration when these forces are measured in the coordinate frame of the head (Merfeld, Zupan et al. 1999). Moreover, different motion stimuli (e.g. translation or tilt) can generate the same sensory signal (Merfeld, Zupan et al. 1999; Angelaki, Shaikh et al. 2004). In order to elicit appropriate responses to translation (e.g. translational vestibulo-ocular reflex (VOR)) and tilt (e.g. postural control) stimuli, the nervous system has to resolve this sensory ambiguity by separating the GIF into neural estimates of gravity and linear acceleration.

In the vestibular system, the semicircular canals measure the angular velocity of the head, while the otolith organs measure the GIF. The semicircular canals keep track of the rate of change of the gravity vector with respect to the head coordinate frame. Merfeld et al. (Glasauer and Merfeld 1997; Merfeld and Zupan 2002) have proposed an internal model which is defined as a neural system that mimics physical principles associated with sensory transduction or movement and can be represented mathematically. This internal model comprises three key components, 1) influence of rotational cues on neural representation of gravity, 2) resolution of GIF into estimation of linear acceleration and gravity, and 3) modeling of the semicircular canals dynamics. The first two components are described by Eqs. 14 and 15:

$$\frac{d\vec{g}}{dt} = -\vec{\omega} \times \vec{g} \quad (14)$$

$$GIF = \vec{g} - \vec{a} \quad (15)$$

where the spatial vectors $\vec{g}$, $\vec{a}$, $\vec{\omega}$ are respectively the estimates for perceived gravity, linear acceleration and angular velocity, all of which are relative to the coordinate frame of the head. The third component is based on a first-order transfer function description of the canal dynamics. The validity of this model has been verified experimentally by measuring the VOR in human subjects following sequential rotation about a vertical axis and tilt about a horizontal axis to produce gravitational illusion (Merfeld, Zupan et al. 1999). Central motion-sensitive

neurons reflecting the computation of such an internal model have been identified in the cerebellar and vestibular nuclei (Angelaki, Shaikh et al. 2004) and in the vestibular cortex, which may also be activated by visual motion inputs (Indovina, Maffei et al. 2005).

The internal model for resolving vestibular sensory ambiguity was proposed in the context of a deterministic state observer (Merfeld, Young et al. 1993; Glasauer and Merfeld 1997; Merfeld and Zupan 2002). When sensory noise and process disturbances are considered, the observer may take the form of a Kalman filter such as that proposed by Park and Kuo (Park and Kuo 2005), or simply a Luenberger observer, to minimize the mean-squared errors of estimation given the sensory inputs. In both cases, the condition of observability (see section on Luenberger observer) is crucial for such an internal model to evolve. This internal model may also play a role in some vestibular control systems, such as postural balance (Kuo 1995). In this event, the explicit cognition of kinematic variables in the adaptation of the internal model is consistent with an indirect adaptive control scheme.

6. Conclusion

In summary, we have provided a broad overview of the study of internal models in sensorimotor integration and their engineering equivalents in adaptive control. It is clear that the internal model paradigm is closely related to adaptive control theory. Rather than relying on computer simulations, performance criteria such as stability, convergence, robustness, observability, controllability and identifiability should be incorporated, whenever possible, in the analysis of internal models as with adaptive control theory. This unified framework should facilitate experimental elucidation of the mechanisms of internal model and the reverse engineering of such neural mechanisms into novel brain-inspired adaptive control paradigms in future.

Acknowledgements

We thank E. Bizzi, N. Hogan, J.-J. Slotine, R. Ajemian and D.M. Wiberg for valuable comments on the manuscript. The authors' work was supported by National Institutes of Health grants HL067966, HL072849 and HL075014.

References

- Abend W, Bizzi E, et al. Human arm trajectory formation. *Brain* 1982;105(Pt 2):331–48. [PubMed: 7082993]
- Andrieu C, Doucet A, et al. Particle Methods for Change Detection, System Identification, and Control. *IEEE Proceedings* 2004;92(3):423–438.
- Angelaki DE, Shaikh AG, et al. Neurons compute internal models of the physical laws of motion. *Nature* 2004;430:560–564. [PubMed: 15282606]
- Anzaldo-Meneses, A. Contemporary trends in nonlinear geometric control theory and its applications. River Edge, NJ: World Scientific; 2002.
- Åström, KJ.; Wittenman, B. Adaptive Control. Addison Wesley; 1995.
- Berniker, M. Mechanical Engineering. Cambridge: Massachusetts Institute of Technology; 2000. A biologically Motivated Paradigm for Heuristic Motor Control in Multiple Context.
- Bertsekas, DP.; Tsitsiklis, J. Neuro-Dynamic Programming. Athena Scientific; 1996.
- Bizzi E, Accornero N, et al. Posture control and trajectory formation during arm movement. *J Neurosci* 1984;4(11):2738–44. [PubMed: 6502202]
- Bizzi, E.; Mussa-Ivaldi, FA. Muscle properties and the control of arm movements. *Visual Cognition and Action*. In: Osherson, DN.; Kosslyn, SM.; Hollerbach, JM., editors. An Invitation to Cognitive Science. 2. Cambridge: MIT Press; 1990. p. 213-242.
- Brown, RG.; Hwang, PYC. Introduction to Random Signals and Applied Kalman Filtering. John Wiley & Sons; 1997.
- Burdet E, Osu R, et al. The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature* 2001;414:446–449. [PubMed: 11719805]

- Camacho, EF.; Bordons, C. Model Predictive Control. Springer; 2004.
- Cheah, CC.; Liu, C., et al. Approximate Jacobian adaptive control for robot manipulators. IEEE International Conference on Robotics and Automation; New Orleans, USA: 2004.
- Cheah, CC.; Liu, C., et al. Adaptive Jacobian Tracking Control of Robots based on Visual Task-space Information. IEEE International Conference on Robotics and Automation; Barcelona, Spain: 2005.
- Doya K. What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw* 1999;12(7–8):961–974. [PubMed: 12662639]
- Doya K, Kimura H, et al. Motor control: Neural models and system theory. *International Journal of Applied Mathematics and Computer Science* 2001;11:101–128.
- Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol* 1954;47:381–391. [PubMed: 13174710]
- Flash T. The Control of Hand Equilibrium Trajectories in Multi-Joint Arm Movements. *Biological Cybernetics* 1987;57:257–274. [PubMed: 3689835]
- Franklin DW, Osu R, et al. Adaptation to Stable and Unstable Dynamics Achieved By Combined Impedance Control and Inverse Dynamics Model. *J Neurophysiol* 2003;90:3270–3282. [PubMed: 14615432]
- Glasauer, S.; Merfeld, DM. Modeling Three Dimensional Vestibular Responses During Complex Motion Stimulation. In: Fetter, M.; Haslwanter, T.; Misslisch, H., editors. Three-dimensional kinematics of eye, head and limb movements. Amsterdam: Harwood academic publisher; 1997. p. 387–398.
- Gomi H, Kawato M. Neural Network Control for a Closed-Loop System Using Feedback-Error-Learning. *Neural Netw* 1993;6(7):933–946.
- Gomi H, Kawato M. Equilibrium-Point Control Hypothesis Examined by Measured Arm Stiffness During Multijoint Movement. *Science* 1996;272:117–120. [PubMed: 8600521]
- Gribble PL, Ostry DJ, et al. Are Complex Control Signals Required for Human Arm Movement? *J Neurophysiol* 1998;79(3):1409–1424. [PubMed: 9497421]
- Hanneton S, Berthoz A, et al. Does the brain use sliding variables for the control of movements? *Biol Cybern* 1997;77(6):381–93. [PubMed: 9433753]
- Hebb, DO. The Organization of Behavior. New York: Wiley; 1949.
- Hodgson AJ, Hogan N. A Model-Independent Definition of Attractor Behavior Applicable to Interactive Tasks. *IEEE Transactions on Systems, Man, and Cybernetics -Part C: Applications and Reviews* 2000;30(1):105–118.
- Hogan N. Adaptive Control of Mechanical Impedance by Coactivation of Antagonist Muscles. *IEEE Transactions on Automatic Control* 1984;29(8):681 – 690.
- Hogan N. Impedance control: an approach to manipulation. Part I. Theory. *J Dyn Sys Meas Con* 1985;107:1–24.
- Hogan, N.; Buerger, SP. Impedance and Interaction Control. In: Kurfess, TR., editor. *Robotics and Automation Handbook*. CRC Press; LLC: 2005. p. 19-1–19–24.
- Huang, S.; Tan, KK., et al. *Applied Predictive Control*. Springer; 2002.
- Imamizu H, Kuroda T, et al. Modular organization of internal models of tools in the human cerebellum. *Proc Natl Acad Sci U S A* 2003;100(9):5461–6. [PubMed: 12704240]
- Imamizu H, Miyauchi S, et al. Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature* 2000;403(6766):192–5. [PubMed: 10646603]
- Indovina I, Maffei V, et al. Representation of visual gravitational motion in the human vestibular cortex. *Science* 2005;308(5720):416–9. [PubMed: 15831760]
- Ioannou, PA.; Sun, J. *Robust Adaptive Control*. Upper Saddle River, NJ: Prentice Hall PTR; 1996.
- Karniel A. Three creatures named ‘forward model’. *Neural Netw* 2002;15(3):305–7. [PubMed: 12125886]
- Kawato M. Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology* 1999;9:718–727. [PubMed: 10607637]
- Kawato M, Furukawa K, et al. A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics* 1987;57:169–185. [PubMed: 3676355]
- Knill, DC.; Richards, W. *Perception as Bayesian Inference*. Cambridge: Cambridge University Press; 1996.

- Kording KP, Wolpert DM. Bayesian integration in sensorimotor learning. *Nature* 2004;427:244–247. [PubMed: 14724638]
- Kuo AD. An optimal control model for analyzing human postural balance. *IEEE Transactions on Biomedical Engineering* 1995;42:87–101. [PubMed: 7851935]
- Kuwata, Y.; How, J. Stable Trajectory Design for Highly Constrained Environments using Receding Horizon Control. *IEEE Proceedings of the American Control Conference*; 2004.
- Li CS, Padoa-Schioppa C, et al. Neuronal correlates of motor performance and motor learning in the primary motor cortex of monkeys adapting to an external force field. *Neuron* 2001;30(2):593–607. [PubMed: 11395017]
- Lohmiller W, Slotine JJE. On Contraction Analysis for Nonlinear Systems. *Automatica* 1998;34(6):683–696.
- Luenberger DG. Observers for Multivariable Systems. *IEEE Transactions on Automatic Control* 1965;11(2):190–197.
- Mareels I. A simple selftuning controller for stably invertible systems. *Systems & Control Letters* 1984;4:5 – 16.
- McClelland, JL.; Rumelhart, DE. *Explorations in parallel distributed processing: a handbook of models, programs, and exercises*. Cambridge, Mass: MIT Press; 1988.
- McIntyre J, Bizzi E. Servo Hypotheses for the Biological Control of Movement. *Journal of Motor Behavior* 1993;25(3):193–202. [PubMed: 12581989]
- Merfeld DM, Young LR, et al. A multidimensional model of the effect of gravity on the spatial orientation of the monkey. *J Vestib Res* 1993;3(2):141–61. [PubMed: 8275250]
- Merfeld DM, Zupan L, et al. Humans use internal models to estimate gravity and linear acceleration. *Nature* 1999;398:615–618. [PubMed: 10217143]
- Merfeld DM, Zupan LH. Neural processing of gravito-inertial cues in humans. III. Modeling tilt and translation responses. *J Neurophysiol* 2002;87(2):819–833. [PubMed: 11826049]
- Miyamura, A.; Kimura, H. *Proc. MTNS-2000*. Perpignan, France: 2000. Synthesis aspects of cerebellum motor control.
- Mussa-Ivaldi FA, Bizzi E. Motor learning through the combination of primitives. *Phil Trans R Soc Lond B* 2000;355:1755–1769. [PubMed: 11205339]
- Mussa-Ivaldi FA, Giszter SF, et al. Linear combinations of primitives in vertebrate motor control. *Proc Natl Acad Sci U S A* 1994;91(16):7534–8. [PubMed: 8052615]
- Mussa-Ivaldi FA, Hogan N, et al. Neural, Mechanical, and Geometric Factors Subservicing Arm Posture in Humans. *The Journal of Neuroscience* 1985;5(10):2732–2743. [PubMed: 4045550]
- Nakanishi, MJ., et al. Learning from demonstration and adaptation of biped locomotion with dynamical movement primitives. *Workshop on Robot Learning by Demonstration, IEEE International Conference on Intelligent Robots and Systems (IROS 2003)*; 2003.
- Neilson, PD.; Neilson, MD., et al. *Approaches to the Study of Motor Control and Learning*. J. J. Summers. Amsterdam: Elsevier Science Publishers; 1992. Adaptive model theory: Application to disorders of motor control; p. 495-548.
- Neilson PD, O'Dwyer NJ, et al. Stochastic prediction in pursuit tracking: an experimental test of adaptive model theory. *Biol Cybern* 1988;58(2):113–22. [PubMed: 3349111]
- Nelson, RJ. *Mechanical Engineering*. Cambridge: Massachusetts Institute of Technology; 1999. *Geometric Control of Quantum Mechanical and Nonlinear Classical Systems*.
- Ogata, K. *Modern Control Engineering*. Prentice-Hall, Inc; 1997.
- Osu R, Hirai S, et al. Random presentation enables subjects to adapt to two opposing forces on the hand. *Nature Neuroscience* 2004;7(2):111–112.
- Padoa-Schioppa C, Li CS, et al. Neuronal correlates of kinematics-to-dynamics transformation in the supplementary motor area. *Neuron* 2002;36(4):751–65. [PubMed: 12441062]
- Padoa-Schioppa C, Li CS, et al. Neuronal activity in the supplementary motor area of monkeys adapting to a new dynamic environment. *J Neurophysiol* 2004;91(1):449–73. [PubMed: 12968016]
- Park S, Kuo AD. An optimal estimator model of multi-sensory processing in human. *On the Convergence of Bio-, Information-, Environmental-, Energy-, Space- and Nano-Technologies*, Pts 1 and 2 *Key Engineering Materials* 2005;277–279:148–154.

- Polderman JW, Mareels I. Two Scale High Gain Adaptive Control. *International Journal of Adaptive Control and Signal Processing* 2004;18:393–402.
- Poon, C.-S.; Tryfonidis, M., et al. Motor learning: a Bayesian decision process?. 30th Soc. for Neurosci. Annual Meeting; New Orleans, LA. 2000.
- Poon CS. Self-tuning optimal regulation of respiratory motor output by Hebbian covariance learning. *Neural Network* 1996;8:1–17.
- Poon, CS. Respiratory models and control. In: Bronzion, JD., editor. *Biomedical Engineering Handbook*. 2. Boca Raton, FL: CRC Press; 2000. p. 161
- Poon CS, Lebowitz HH, et al. Negative-impedance ventilation and pressure support ventilation: a comparative study. *Respir Physiol* 1997;108(2):117–27. [PubMed: 9232685]
- Poon CS, Siniatia MS. Plasticity of cardiorespiratory neural processing: classification and computational functions. *Respir Physiol* 2000;122(2–3):83–109. [PubMed: 10967337]
- Poon CS, Ward SA. A device to provide respiratory-mechanical unloading. *IEEE Trans Biomed Eng* 1986;33(3):361–5. [PubMed: 3082741]
- Prinz AA, Abbott LF, et al. The dynamic clamp comes of age. *Trends Neurosci* 2004;27(4):218–24. [PubMed: 15046881]
- Rancourt D, Hogan N. Stability in force-production tasks. *J Mot Behav* 2001;33(2):193–204. [PubMed: 11404214]
- Rao, RP.; Olshausen, BA., et al. *Perception and neural function*. Cambridge, MA: MIT Press; 2002. Probabilistic models of the brain.
- Redish AD. Addiction as a Computational Process Gone Awry. *Science* 2004;306(5703):1944–1947. [PubMed: 15591205]
- Schaal, S. Dynamic movement primitives - A framework for motor control in humans and humanoid robots. *The International Symposium on Adaptive Motion of Animals and Machines*; 2003.
- Schaal, S.; Peters, J., et al. *Learning Movement Primitives*. *International Symposium on Robotics Research (ISRR2003)*; 2004.
- Schultz W, Dayan P, et al. A neural substrate of prediction and reward. *Science* 1997;275(5306):1593–9. [PubMed: 9054347]
- Shadmehr R, Mussa-Ivaldi FA. Adaptive representation of dynamics during learning of a motor task. *J Neurosci* 1994;14(5 Pt 2):3208–24. [PubMed: 8182467]
- Shidara M, Kawano K, et al. Inverse-dynamics model eye movement control by Purkinje cells in the cerebellum. *Nature* 1993;365(6441):50–2. [PubMed: 8361536]
- Slotine, J-J.; Li, W. Adaptive manipulator control a case study. *IEEE International Conference on Robotics and Automation*; 1987.
- Slotine, J-JE.; Li, W. *Applied Nonlinear Control*. Prentice Hall: 1991.
- Slotine, JJ.; Li, W. On the Adaptive Control of Robot Manipulators. *A.S.M.E. Winter Annual Meeting*; Anaheim, CA. 1986b.
- Slotine JJ, Lohmiller W. Modularity, evolution, and the binding problem: a view from stability theory. *Neural Netw* 2001;14(2):137–45. [PubMed: 11316230]
- Slotine JJE. Modular Stability Tools for Distributed Computation and Control. *Int J Adaptive Control and Signal Processing* 2003;17(6)
- Slotine JJE, Coetsee JA. Adaptive Sliding Controller Synthesis for Nonlinear Systems. *Int J Control* 1986a;43(4)
- Slotine JJE, Li W. On The Adaptive Control of Robot Manipulators. *Int J Robotics Research* 1987;6(3)
- Sober SJ, Sabes PN. Multisensory integration during motor planning. *J Neurosci* 2003;23(18):6982–92. [PubMed: 12904459]
- Suri RE. Anticipatory responses of dopamine neurons and cortical neurons reproduced by internal model. *Exp Brain Res* 2001;140(2):234–40. [PubMed: 11521155]
- Sutton, RS.; Barton, AG. *Reinforcement Learning - An Introduction*. Cambridge, MA: The MIT Press; 1998.
- Thach WT, Goodkin HP, et al. The Cerebellum and the Adaptive Coordination of Movement. *Annual Review of Neuroscience* 1992;15:403–442.

- Tryfonidis, M. Department of Aeronautics and Astronautics. Cambridge, MA: Massachusetts Institute of Technology; 1999. Robust Adaptive Control Modeling of Human Arm Movements subject to Altered Gravity and Mechanical Loads.
- Vercher JL, Sares F, et al. Role of sensory information in updating internal models of the effector during arm tracking. *Prog Brain Res* 2003;142:203–22. [PubMed: 12693263]
- Wiberg DM, Powell TD, et al. An On-Line Parameter Estimator for Quick Convergence and Time-Varying Linear Systems. *IEEE Transactions on Automatic Control* 2000;45(10):1854–1863.
- Wolpert DM, Ghahramani Z, et al. An internal model for sensorimotor integration. *Science* 1995;269(5232):1880–2. [PubMed: 7569931]
- Wolpert DM, Kawato M. Multiple paired forward and inverse models for motor control. *Neural Network* 1998;11:1317–1329.
- Wolpert DM, Miall RC. Forward Models for Physiological Motor Control. *Neural Netw* 1996;9(8):1265–1279. [PubMed: 12662535]
- Young, DL. Mechanical Engineering. Cambridge, MA: M.I.T; 1997. Hebbian Covariance Learning and Self-Tuning Optimal Control.
- Young DL, Poon CS. Hebbian covariance learning. A nexus for respiratory variability, memory, and optimization? *Adv Exp Med Biol* 1998;450:73–83. [PubMed: 10026966]
- Young DL, Poon CS. A Hebbian feedback covariance learning paradigm for self-tuning optimal control. *IEEE Transactions on Systems, Man, and Cybernetics - Part B* 2001;31:173–186.
- Zhao Y, Slotine J. Discrete nonlinear observers for inertial navigation. *Systems & Control Letters* 2005;54(9):887–898.

Fig. 1.

Fig. 1a Direct Adaptive Control

The controller is adapted directly from the input/output signals.

Fig. 1b Indirect Adaptive Control

The controller is adapted by first estimating the plant parameters and/or state variables.

Fig. 2. Full-order Luenberger Observer

To identify unmeasurable state variables, an observer in the same form as the plant assumed is used. The difference between the plant and observer outputs is weighted and sent to the observer to modify its output. The error weighting is chosen such that the output error will converge to zero provided the condition of observability is satisfied.

Fig. 3. Kalman Filter Recursive Loop

The Kalman filter recursively estimates the state variables from noisy measurements. It assumes linear system dynamics and Gaussian measurement noise. The algorithm computes an optimal update gain which minimizes the Bayesian mean-square errors for unbiased estimates.

Fig. 4. High-gain Adaptive Control (HGAC)

A proportional gain controller can stabilize an unstable plant by increasing the feedback gain, K , until the feedback error tends to zero, provided the SPR conditions are satisfied.

Fig. 5. Gain Scheduling

A set of controllers are pre-designed offline for varying operating conditions of the plant. The actual controller at every instant is chosen online based on the current operation point.

Fig. 6. MRAC Architecture

A reference model supervises the desired behavior of the closed-loop system. The outputs of the reference model and the system are compared and this error signal is used to adapt the feedforward and feedback controllers until the system converges to the reference model.

Fig. 7. Adaptive Sliding Control

The algorithm comprises a fixed feedback and an adaptive feedforward loop. The sliding variable, s , drives the feedback controller and also adapts the feedforward controller. Unknown system parameters are estimated online. Convergence of tracking error and estimation error is guaranteed by Lyapunov analysis.

Fig. 8. Feedback Error Learning

The algorithm is similar to adaptive sliding control except the output of the feedback controller is used to adapt the feedforward controller which forms the inverse internal model of the system.

Fig. 9. Multi-model Adaptive Control

A set of reference models and controllers corresponding to differing operation points is given *a priori*. At every instant, the outputs of the reference models are compared to the actual output of the plant. An optimal reference model that best fits the plant is then selected by a switching mechanism to compute the control input.

Fig. 10. Model Predictive Control

At every instant the control signal is computed by optimizing a cost function over a finite number of future time steps (planning horizon). Only a subset of calculated control input is executed each time (execution horizon).

Fig. 11. Reinforcement Learning

The value function specifies the expected reward for a specific action defined by the policy executed at a specific state. Difference between the expected and actual rewards updates the value function and the policy in a way to maximize the returned reward.

Fig. 12. Hebbian Feedback Covariance Control Paradigm

The controller gain, C , is adapted based on the covariance between the input (y) and output (u) signals. Dotted line indicates the original control scheme for static system. To account for system dynamics, a filter is added to transform the plant output, y into an intermediate variable, z , as given by Eq. 12.

Fig. 13. Self-tuning Model of Respiratory Control

$P_{i\text{CO}_2}$ is partial pressure of CO_2 in inspired air, P_c is chemoreceptor activity, $\dot{V}_{\text{CO}_2}$ is metabolic CO_2 production rate and $\dot{V}_E$ is ventilation rate. Self-tuning control is achieved through synaptic plasticity in respiratory neuron (RN) which drives a motor neuron (MN). This respiratory control model resembles the control system in Fig. 12.

Table 1

Equivalence of adaptive control theory and varying hypotheses of motor learning

Hypothesis	Engineering Equivalent	Direct or Indirect?	Proof of convergence?	Comments
Equilibrium Point	Gain scheduling	Direct	By simulation	Fast reacting but can be costly and non-robust
Feedback Error Learning	Adaptive sliding control	Direct	Convergence guaranteed by Lyapunov analysis	Require linearly parameterizable systems
Combined Impedance Control and Inverse Model	High-gain adaptive control and a second adaptive control (e.g. adaptive sliding control)	HGAC is direct and the second controller can be either direct or indirect	HGAC requires SPR condition, convergence for the combined control is unclear	The combined controller should out-perform individual controller
Modular paired forward and inverse models	Multi-model adaptive control	Indirect with state estimation in the forward model	By simulation	Flexible with biological relevance