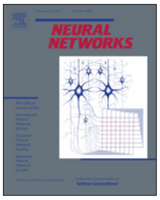
Neural Networks ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet



2013 Special Issue

Adaptive filters and internal models: Multilevel description of cerebellar function

John Porrill[a], Paul Dean[a,*], Sean R. Anderson[b]

[a] *Department of Psychology, Sheffield University, Western Bank, Sheffield, S10 2TP, UK*
[b] *Department of Automatic Control and Systems Engineering, Sheffield University, Sheffield, S1 3JD, UK*

ARTICLE INFO

Keywords:
Cerebellum
Internal model
Adaptive control

ABSTRACT

Cerebellar function is increasingly discussed in terms of engineering schemes for motor control and signal processing that involve internal models. To address the relation between the cerebellum and internal models, we adopt the chip metaphor that has been used to represent the combination of a homogeneous cerebellar cortical microcircuit with individual microzones having unique external connections. This metaphor indicates that identifying the function of a particular cerebellar chip requires knowledge of both the general microcircuit algorithm and the chip's individual connections.

Here we use a popular candidate algorithm as embodied in the adaptive filter, which learns to decorrelate its inputs from a reference ('teaching', 'error') signal. This algorithm is computationally powerful enough to be used in a very wide variety of engineering applications. However, the crucial issue is whether the external connectivity required by such applications can be implemented biologically.

We argue that some applications appear to be in principle biologically implausible: these include the Smith predictor and Kalman filter (for state estimation), and the feedback–error–learning scheme for adaptive inverse control. However, even for plausible schemes, such as forward models for noise cancellation and novelty-detection, and the recurrent architecture for adaptive inverse control, there is unlikely to be a simple mapping between microzone function and internal model structure.

This initial analysis suggests that cerebellar involvement in particular behaviours is therefore unlikely to have a neat classification into categories such as 'forward model'. It is more likely that cerebellar microzones learn a task-specific adaptive-filter operation which combines a number of signal-processing roles.

1. Introduction: the chip metaphor

Recent reviews indicate that the possible role of the cerebellum in the formation of internal models is a topic of growing interest (Bastian, 2011; Cerminara & Apps, 2011; Ebner, Hewitt, & Popa, 2011; Imamizu, 2010; Medina, 2011; Shmuelof & Krakauer, 2011). Although the term 'internal model' can be used very generally in this context to refer to any neural representation of a dynamic system (Wolpert, Ghahramani, & Jordan, 1995), many of its most important conceptual features can be captured by a simple example (Fig. 1).

Motor commands (as expressed by motoneurons) act on muscles which move a part of the body. The mechanical properties of the muscles and body part (for convenience referred to as the motor plant) ensure that the dynamics of the movement differ from those of the command (Fig. 1(A)). The circuit shown in Fig. 1(B) allows a model of this plant to be learnt, by sending a copy of the motor commands to an adaptive element (highlighted in red throughout). The output of this element is compared with sensory feedback from the actual movement, and the discrepancy between the two used as a signal to 'train' the adaptive element (a form of supervised learning). As training proceeds the discrepancy decreases, meaning that the dynamics of the adaptive element resemble more closely the dynamics of the motor plant. In this way a model of the motor plant is learnt.

The model shown in Fig. 1 (usually referred to as a 'forward' model, as explained below) has many uses, including the prediction of the sensory effects of movement. Such a prediction can, for example, help distinguish the sensory signals produced by one's own movements from those arising from external events—the classical reafference problem (e.g. Cullen, 2004), and further uses are discussed in Section 3. Hence proposals that the cerebellum takes part in the formation of internal models seek to provide a crucial link between cerebellar function and proven sensorimotor competences (Blakemore, Frith, & Wolpert, 1999, 2001; Imamizu, Kuroda, Miyauchi, Yoshioka, & Kawato, 2003; Kawato, 1995, 1996, 1999, 2008; Miall, Christensen, Cain, & Stanley, 2007; Miall & Reckess, 2002; Miall, Weir, Wolpert, & Stein, 1993; Miall & Wolpert, 1996; Wolpert, 1997; Wolpert & Ghahramani, 2000; Wolpert et al., 1995; Wolpert & Kawato, 1998; Wolpert, Miall, & Kawato, 1998).

---

* Corresponding author.
E-mail address: p.dean@sheffield.ac.uk (P. Dean).

0893-6080/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2012.12.005

Please cite this article in press as: Porrill, J., et al. Adaptive filters and internal models: Multilevel description of cerebellar function. Neural Networks (2013), doi:10.1016/j.neunet.2012.12.005

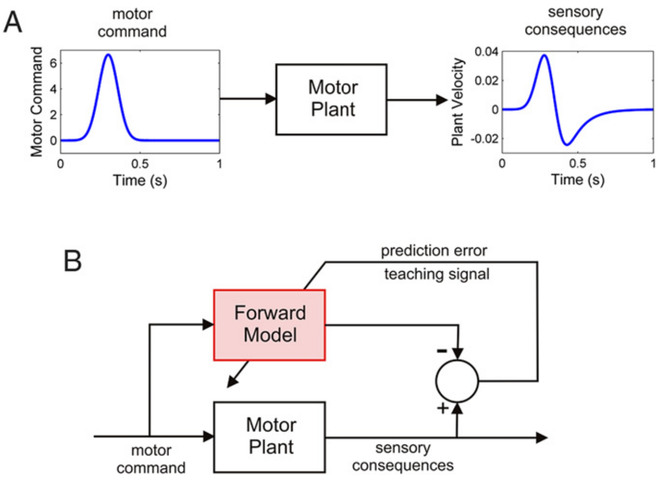


Fig. 1. Dynamic response and forward model of a simple viscoelastic motor plant. A: The motor command from the motoneurons acts on muscles, which move some part of the body. The mechanics of the muscles plus body part (= 'motor plant') mean that the temporal trajectory of the movement differs from that of the motor command of the profile of the plant. The example shows the velocity response of a simple viscoelastic plant to a Gaussian motor command. (For convenience, the dynamics of sensory transduction are neglected, so the sensory measurement of the movement introduces no distortions.) B: A circuit for learning a forward model of the motor plant dynamics. The forward model is the adaptive element (highlighted in red; this convention is also applied in later figures). It can be learnt using sensory error (that is, the difference between the predicted and actual sensory consequences of the motor command) as teaching signal. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

To evaluate how far this proposed link is supported by current evidence, the present article first outlines the popular 'chip' metaphor for cerebellar organisation, which requires cerebellar functions to be modelled at both microcircuit and external-connectivity levels.

1.1. The 'chip' metaphor

The arrangement of neurons and their connections within cerebellar cortex is broadly similar over the entire cortical mantle, whereas each individual region of the cerebellar cortex has a unique pattern of connections with external neural structures. This combination has long been recognised:

> "The cerebellar tissues have quite a uniform histological structure. Their role in the actual motor control, however, varies from region to region, depending upon what subcortical structures they are connected with, as pointed out by Herrick (1924)" (Ito, 1970, p. 162);

and has given rise to what might be termed the 'chip' metaphor of cerebellar organisation

> "Cerebellar microcomplexes are connected to various systems of the brain and so play diverse roles in central nervous system functions. This situation would be similar to that of a computer chip which can be used for a great many purposes." (Ito, 1997, p. 475).

This metaphor is illustrated in Fig. 2, which shows in schematic form a functional sub-region of cerebellar cortex represented by an identical internal structure and idiosyncratic inputs and outputs. The important implication of the chip metaphor is that the function of any particular cerebellar sub-region depends on *both* the signal-processing capacities of the generic chip, *and* the particular architecture in which it is embedded.

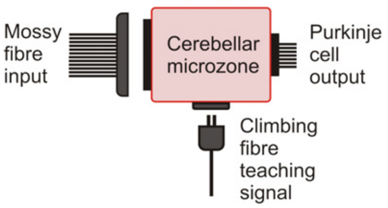


Fig. 2. The cerebellar 'chip' metaphor. Each cerebellar microzone has a similar internal organisation, but its own idiosyncratic set of connections, two inputs and one output. The climbing fibre and output connections are unique to a microzone: some of a microzone's mossy fibre inputs may be shared with other microzones. The climbing fibre teaching signal specifies the learning goals for the chip, hence it is this connectivity which is basic to defining individual microzones. The Purkinje cell output must then be connected to a target region in the deep cerebellar or vestibular nuclei which contributes to achieving this goal, and for which the learning procedure hardwired into the chip is stable and convergent. This provides a strong constraint on the output connectivity. The mossy fibre input connections are the least constrained. They can be regarded as a wide 'bus' of possibly relevant sensory and motor signals, from which those signals actually relevant to the task will be chosen by the learning procedure.

The relevance of the chip metaphor for evaluating internal-model hypotheses of cerebellar function can be illustrated by the 'inverse-model' circuit shown in Fig. 3. The need for an inverse model of the motor plant arises because of the 'distorting' effects of plant dynamics on the motor command, as shown in Fig. 1(A). Motor commands that specify a desired trajectory for a part of the body must therefore be converted into a form that compensates for the characteristics of the plant. This can be achieved by passing the command, not directly to the plant itself, but indirectly through an inverse model of the plant (Fig. 3(A)). As with the 'forward' plant model (terminology emphasising the contrast with the inverse plant model) such a model needs to be learnt, and a possible circuit for achieving this is shown in Fig. 3(B). Although, as will be argued later, the circuit shown in Fig. 3(B) is too simple to be biologically realistic, it illustrates an important point about the difference between an adaptive element and the circuit of which it is a component. Comparison of Figs. 1 and 3 shows how the same adaptive element can learn either a forward model, or an inverse model, depending on the details of the external wiring. This is exactly the point captured by the cerebellar 'chip' metaphor of Fig. 2.

Evaluating internal-model hypotheses of cerebellar function therefore entails evaluating both the microcircuit model and the way it is wired into any particular system-level architecture. The particular microcircuit model chosen here is the adaptive filter, and this is briefly described in Section 2, and its general computational suitability for internal model formation explained. Particular internal-model architectures are then assessed in two stages. The first asks how they are biologically plausible—how far the signals they require could be in principle provided biologically (Section 3). The second stage considers the problems that arise when the plausible architectures have to be mapped in practice onto complex neural circuits (Section 4). The final section addresses the implications of the internal-model hypothesis for the future understanding of cerebellar functions (Section 5).

2. Microcircuit level: inside the chip

The repeating nature of the cerebellar microcircuit suggests that there is a generic 'cerebellar algorithm', and hypotheses about its computational capability (Albus, 1971; Marr, 1969) appeared soon after the microcircuit itself was first described (Eccles, Ito, & Szentágothai, 1967). The Marr–Albus framework was further developed by Fujita (1982), who proposed that the cerebellar circuit acts as an adaptive filter in which the mossy fibre inputs to the cerebellum convey dynamic time-varying signals rather than the static spatial patterns associated with the original Marr–Albus formulations (further details in Section 2.4). Since it appears that

Please cite this article in press as: Porrill, J., et al. Adaptive filters and internal models: Multilevel description of cerebellar function. Neural Networks (2013), doi:10.1016/j.neunet.2012.12.005