



## Review

**Cite this article:** Morimoto J, Kawato M. 2015 Creating the brain and interacting with the brain: an integrated approach to understanding the brain. *J. R. Soc. Interface* **12**: 20141250.  
<http://dx.doi.org/10.1098/rsif.2014.1250>

Received: 11 November 2014  
Accepted: 17 December 2014

### Subject Areas:

computational biology

### Keywords:

computational neuroscience, humanoid robot, exoskeleton robot, neurofeedback, motor learning

### Authors for correspondence:

Jun Morimoto  
e-mail: [xmorimo@atr.jp](mailto:xmorimo@atr.jp)  
Mitsuo Kawato  
e-mail: [kawato@atr.jp](mailto:kawato@atr.jp)

# Creating the brain and interacting with the brain: an integrated approach to understanding the brain

Jun Morimoto<sup>1</sup> and Mitsuo Kawato<sup>2</sup>

<sup>1</sup>Department of Brain Robot Interface, ATR Computational Neuroscience Labs, Kyoto, Japan

<sup>2</sup>ATR Brain Information Communication research Laboratory Group, Kyoto, Japan

In the past two decades, brain science and robotics have made gigantic advances in their own fields, and their interactions have generated several interdisciplinary research fields. First, in the ‘understanding the brain by creating the brain’ approach, computational neuroscience models have been applied to many robotics problems. Second, such brain-motivated fields as cognitive robotics and developmental robotics have emerged as interdisciplinary areas among robotics, neuroscience and cognitive science with special emphasis on humanoid robots. Third, in brain–machine interface research, a brain and a robot are mutually connected within a closed loop. In this paper, we review the theoretical backgrounds of these three interdisciplinary fields and their recent progress. Then, we introduce recent efforts to reintegrate these research fields into a coherent perspective and propose a new direction that integrates brain science and robotics where the decoding of information from the brain, robot control based on the decoded information and multimodal feedback to the brain from the robot are carried out in real time and in a closed loop.

## 1. Introduction

Not only scientists but also robotics engineers have been attracted to the elucidation of information processing in the brain. The human brain, which is a complex dynamical system with deep hierarchy and huge degrees of freedom, possesses  $10^{11}$  neurons [1]. Big projects that have recently started in the USA [2] and the EU [3] ultimately aim to record from all the neurons in the brain, experimentally controlling them and simulating their working by computer. Even if we could record the neural activities of all the neurons, experimentally control their activities and simulate neural circuit behaviours, we would still be far from understanding the brain’s information processing unless we understand the computational principles that underlie the problems solved by it [4]. Robotics is a research field in engineering that could be integrated with brain science to elucidate the computational principles relevant to brain information processing [5].

The interdisciplinary fields between brain science and robotics can be classified into the following three types: (1) computational neuroscience in a narrow sense that models the various levels of events and information processing in the brain. This includes representative neural network models [6–11]; (2) brain-motivated robotics, where researchers build an artificial system that reproduces the brain’s information processing and generally explore brain-like computations [12]. Different subfields have different emphasis on machine-learning, nonlinear dynamics, neuroscience, cognitive science [13], developmental psychology [14] and so on. But all of them possess some relevance to brain science. Imitation learning, which is partially inspired by a neuroscience study [15], has become one of the most popular approaches in brain-motivated robotics and is now widely studied [12,13,16–18]. In recent years, developmental robotics have also been intensively studied [14,19]; (3) a brain–machine interface (BMI), which is sometimes called a brain–computer interface (BCI), connects the brain with such external devices as spellers [20] or prosthetics [21]. A BMI can be defined as an interface that

connects the brain and a computer/robot based on some computational neuroscience theories [22–26]. Typically, a computational model is used to extract necessary information from the brain based on signals, including neural firing, local field potentials, electroencephalography, near infrared spectroscopy and functional magnetic resonance imaging (fMRI) signals. The ‘decoded’ information [27] from the brain is fed to a robotic arm, for example, and visual and/or tactile feedback to the brain closes the loop between the brain and the robot. BMIs are used not only to compensate and substitute for lost or damaged sensory-motor functions including communications but also to recover and repair such functions [22–26,28–33]. Furthermore, there are multiple ongoing projects that are aiming to recover the biped walking functions of spinal-cord injury and stroke patients [29,31,34,35].

The application of computational neuroscience models to robots opened up an approach called ‘understanding the brain by creating the brain’ [5,36], which was emphasized in the Japanese promotion of neuroscience about two decades ago [37]. To fully validate the computational feasibility and the efficacy of some theories, we need to apply them to real-world problems; we need an artificial brain and body. The former might be realized as computer software, but the latter necessitates robots as physical substitutions for the body. Actual sensory-motor control and decision-making problems are obliged to deal with a huge amount of d.f., the low signal-to-noise ratios of sensory signals and motor outputs, real-time constraints on computation and feedback, and complicated nonlinear events that are difficult to model mathematically, such as contact dynamics and friction. Using robots forces us to deal with all these difficulties, while computer simulations tend to overlook such difficult boundary conditions and underestimate the complications of the computational problems solved by the brain. Computational neuroscience studies and brain-motivated robotics studies are investigating how the brain can acquire action policies to control its own body, which has a gigantic number of degrees of freedom, while receiving environmental information from a huge number of sensory inputs. Note that the term ‘policy’, which was originally borrowed from reinforcement-learning literature [11], corresponds to ‘control law’ or ‘control rule’ in control engineering studies. These computational neuroscience and brain-motivated robotics can be considered ‘creating the brain’ approaches. On the other hand, BMI studies focused on decoding brain activities to control external devices or extracting brain activities to feed the information back to subjects so that they can modulate their brain activities. BMI research paradigms can be considered ‘interacting with the brain’ approaches.

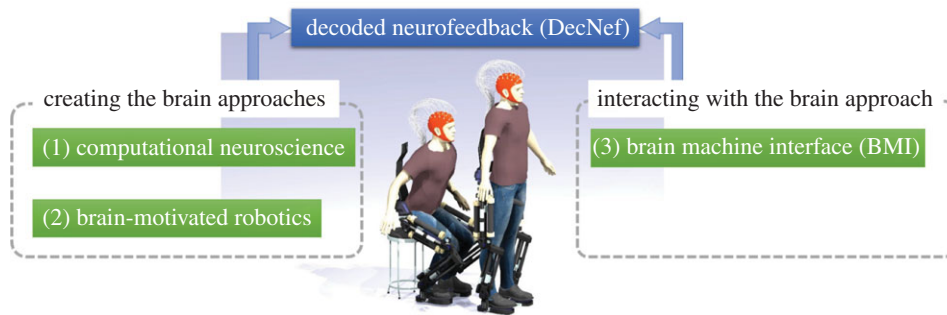
Recent advances suggest that these previously divergent fields between robotics and brain science can be reintegrated into novel methodologies with higher impacts on science and technology. For example, an electroencephalogram (EEG) BMI-based neurorehabilitation system for stroke patients has been developed [33], where the BMI rehabilitation approach is regarded as an integrated system of the above brain-motivated robotics and BMI elements. The decoded neurofeedback (DecNef) method is another concrete example of the integration of the previously divergent way of understanding brains, because it integrates computational neuroscience and BMI disciplines to experimentally induce spatial voxel patterns in a limited brain region, facial preference and colour. DecNef has already realized the perceptual learning of specific orientation

gratings [38], the associative learning for the manipulation of facial preference [39] and the creation of the phenomenal consciousness of colour [40].

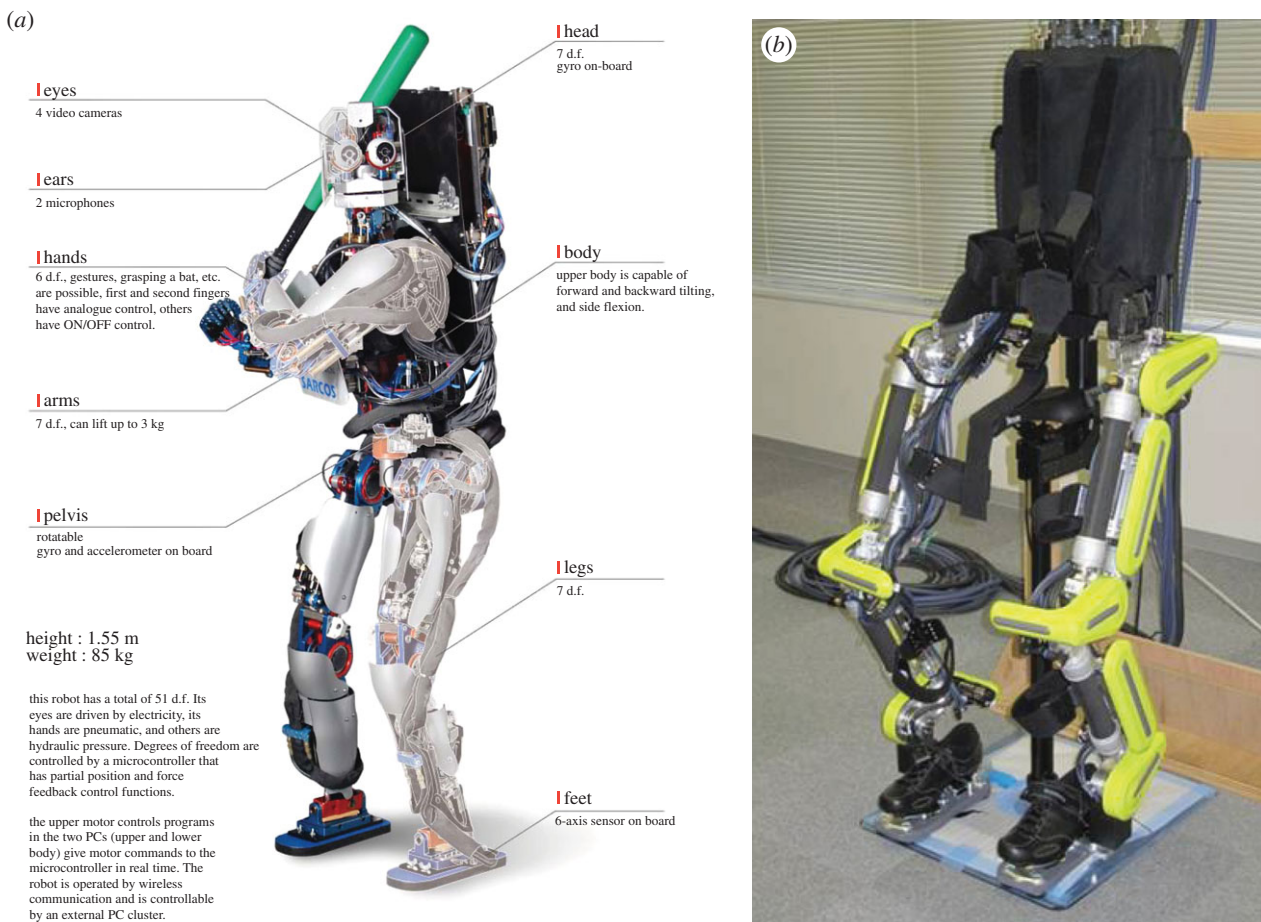
The main purpose of this review is to motivate future research directions to reintegrate the above three research fields that straddle robotics and brain science. These three interdisciplinary research fields have made gigantic advances over the past two decades, mainly in parallel. However, more recently, we see more and more interesting research examples that reunite these previously divergent directions, as exemplified above. We argue here that the future of computational neuroscience in a broader sense is the integration of all three disciplines: computational neuroscience, brain-motivated robotics and BMI. In this review, we introduce efforts to bridge and combine them by postulating a new research direction that fully integrates them. Within this paradigm, a brain and a robot are bi-directionally coupled by decoding [27] and physical interaction. A human volunteer wears a whole-body exoskeleton robot. Her/his brain activity is measured and decoding is carried out in real time. The decoded brain information is used to influence the robot control algorithms to realize brain-to-robot information transmission. Because the robot is attached to a human body, robot motion generates multi-modal sensory feedback to the brain by implementing robot-to-brain information transmission. This system could compensate for lost functions, perform therapeutic treatments of disorders or induce significant spatio-temporal neural activity and brain plasticity. A schematic diagram of this unified direction is depicted in figure 1. We provide several theoretical arguments why this experimental paradigm is a powerful approach to reunite the three interdisciplinary research fields between brain science and robotics. Such reunification provides better understanding of the brain using the combined brain-related methodologies and technologies. Simultaneously, future systems developed within this reunification will be very useful in BMI-based neurorehabilitation or revolutionary therapeutics for psychiatric disorders. We then discuss the current research trends in which computational theory is used for brain decoding in BMI and an exoskeleton robot is used for neurofeedback to the brain.

We first review our proposed approach by introducing our computational neuroscience models and brain-motivated humanoid-motor-learning studies and then introduce our BMI exoskeleton robot studies. We discuss our novel unified view of creating the brain and interacting with it approaches by introducing our new humanoid exoskeleton studies and our recently developed DecNef method. For the brain-motivated humanoid-motor-learning studies, we explored the creating the brain approach by developing a series of humanoid robots [36,41], including CB-i, a human-sized humanoid robot (figure 2a) [42]. As a new generation of humanoids, we are developing exoskeleton robots that can physically interact with the human body to assist user movements (figure 2b) [43,44]. An important technical element of DecNef is the real-time neurofeedback of decoded information to the brain. Therefore, as we explained above, a humanoid exoskeleton can be used as an output device to provide a rich set of feedback signals to the body and ultimately to the brain.

The rest of this paper is organized as follows. In §2, we introduce our computational neuroscience models to understand the functional brain mechanisms. We explore modular and hierarchical architectures as possible approaches to deal



**Figure 1.** We categorize brain-related theories, models and methodologies into three disciplines: (1) computational neuroscience, (2) brain-motivated robotics and (3) BMI. Computational neuroscience and brain-motivated robotics are ‘creating the brain’ approaches. On the other hand, BMI methods are ‘interacting with the brain’ approaches. The future of computational neuroscience in a broader sense is the integration of all three disciplines. Integration could take the concrete form of a human volunteer wearing a humanoid exoskeleton controlled by brain activity measured, and rich information feedback to the volunteer’s body and brain while covering a wide range of sensory modalities based on DecNef [38] methodology. This system might compensate for lost functions, perform therapeutic treatments of disorders or induce significant spatio-temporal neural activity and brain plasticity. (Online version in colour.)



**Figure 2.** (a) Human-sized humanoid robot, CB-i [42]: height: 1.55 m, total weight: 85 kg. The robot is equipped with 51 d.f.:  $2 \times 7$  d.f. legs,  $2 \times 7$  d.f. arms,  $2 \times 2$  d.f. eyes, 3 d.f. neck/head, 1-d.f. mouth, 3 d.f. torso and  $2 \times 6$  d.f. hands. It has similar configurations, range of motion, power and strength as a human body to reproduce better natural human-like movements, particularly for locomotion and object manipulation. (b) Exoskeleton robot (XoR) [43]. Height: 1.5 m, weight: 30 kg. XoR has 10 d.f. and six active joints, each of which uses a hybrid actuator composed of an air muscle and an electric motor. It is designed to assist lower limb movements of human users. (Online version in colour.)

with high-dimensional systems. In §3, we review our previous research activities using humanoid robots to understand the brain’s motor learning mechanisms. In particular, in §3.2, we discuss possible approaches to understand how the brain can acquire behavioural policies using extremely high-dimensional brain activities while simultaneously receiving a huge amount of sensory inputs. In §4, we introduce our newly developed BMI exoskeleton robot and investigate how its movements can affect the subject’s brain activities. Finally, in §5, we

discuss the possible applications of an exoskeleton robot in the DecNef framework.

## 2. Creating a brain by computational neuroscience models

Here, we introduce computational neuroscience models to cope with high-dimensional state-space problems. Policy learning in high-dimensional state space is extremely difficult

from a computational modelling point of view. To derive an optimal control policy for nonlinear dynamics, dynamic programming [45] and reinforcement learning [11] are used as the standard policy-learning algorithms. In them, a tremendous number of learning iterations are required to find a policy in high-dimensional state space; this difficulty is called the ‘curse of dimensionality’.

In §§2.1 and 2.2, we discuss possible approaches to understand how the brain copes with this policy-learning problem in high-dimensional state space.

### 2.1. Modular learning approaches

One possible solution for handling large-scale problems is to divide and conquer. As a concrete approach, we introduce the modular selection and identification control (MOSAIC) model, which was originally proposed for supervised learning [46,47] and then extended to reinforcement learning [48]. The most advanced form of the reinforcement MOSAIC [49] has three different sets of expert networks, each of which is specialized for each subproblem. One of these network sets is composed of the predictors of state transitions, which are formally called internal forward models. An internal forward model receives the information of the current state and the current motor command and then predicts the next time-step state. The two other sets approximate actual rewards and reinforcement learning controllers that compute motor commands for optimal control. The responsibility signal for the internal forward model network is determined by the prediction error in the dynamics, and the responsibility signal for the other two networks is determined by the approximation error in the reward and in the reward-prediction error; these signals are used to weight the outputs from the experts. The signals also gate the learning of each expert network by regulating the learning rate.

Switching and weighting by approximation and prediction errors are key features of MOSAIC. The medial prefrontal cortical areas might play an important role in detecting environmental changes by accumulating prediction error. The lesions of monkey medial prefrontal cortex affect the length of the accumulated reward-prediction error in contextual switching behaviour [50]. The responsibility signal could be interpreted as the posterior probability of the Bayesian inference, assuming multiple linear systems have Gaussian noise. In a human imaging study [51], ventromedial prefrontal activities in a stochastic reversal task were correlated with the Bayesian update errors of probability, which are the differences between the posterior and prior probabilities. Circumstantial support of hierarchy and modularity in the context of reinforcement learning was mainly obtained from neuroimaging and lesion studies. The activities of many brain areas, including those of the prefrontal cortex, the orbitofrontal cortex, the premotor cortex, the supplementary motor area, the cerebellum, and the basal ganglia, are correlated with such important variables in reinforcement learning as accumulated reward, learned behaviour, short-term reward and behavioural learning, even in a simple Markov decision process with monetary reward [52]. Furthermore, although activity in the human putamen is correlated with reward prediction that depends on selected actions, which is consistent with a study in monkeys [53], activity in the caudate is correlated more closely with reward-prediction error [54–56]. Topographic representations with different parameters of the discount factor for reward prediction were found in the medial prefrontal and insular cortices, and

topographic representations for reward-prediction error were found in the basal ganglia [57].

As described above, exploring how our proposed model can be implemented as a neural system is crucial for computational neuroscience to understand how the brain deals with real-world control problems.

### 2.2. Hierarchical learning approaches

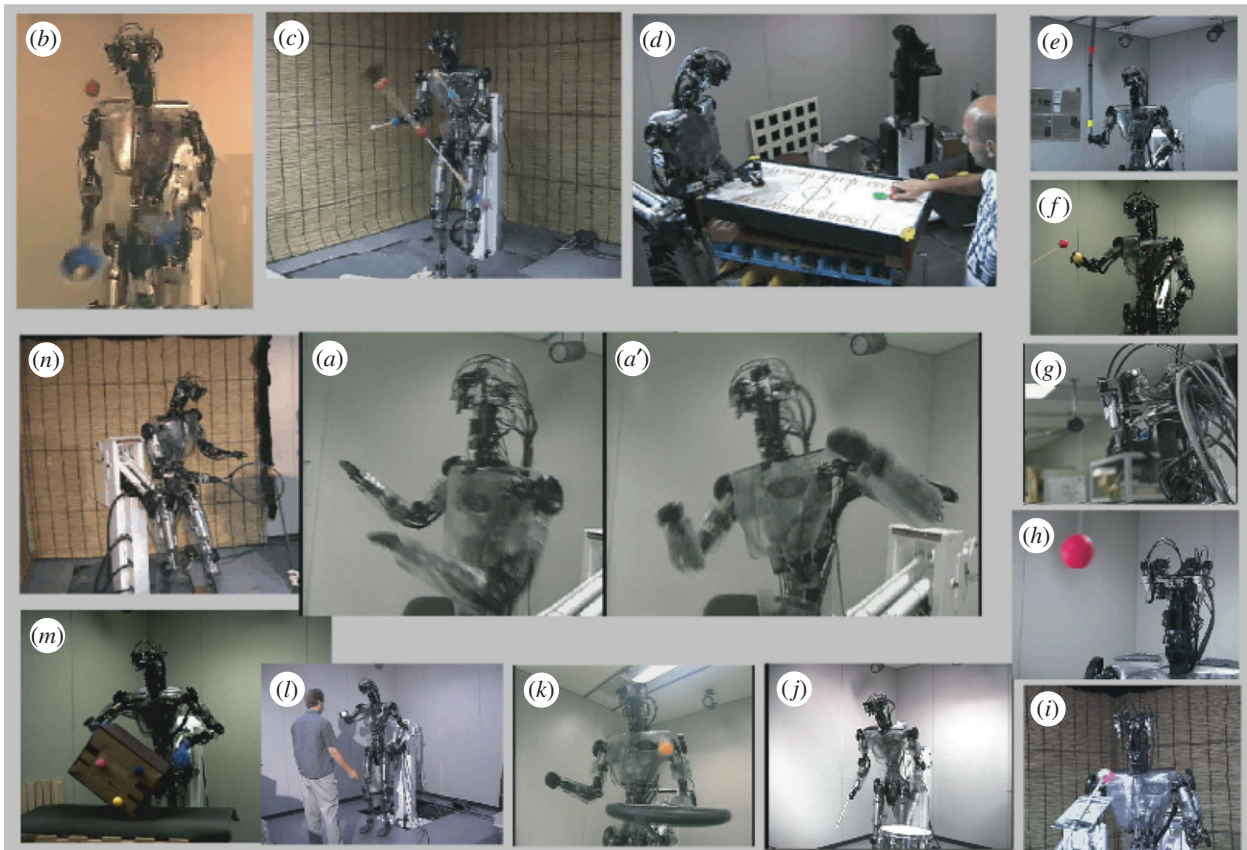
The other common strategy to cope with the high-dimensional policy-learning problem is to add hierarchy to the algorithms [58,59]. In the upper part of such a hierarchy, an approximate optimal solution could be found in a reasonable time because the state space that describes a body and an external world is coarsely represented by a reasonable number of discrete regions, cells or neurons (e.g. up to 1 million neurons, a reasonable number for one function). In the lower layer of this hierarchy, the state space is finely represented and the approximate solution that is obtained in the upper layer can constrain the possible range of the exploration space. Thus, the problem could be solved within a reasonable amount of time and with the required accuracy. In most studies, coarse representations have been specified in the upper hierarchy or used to predetermine the motor primitives and/or the subgoals. This recourse is similar to assuming the existence of a homunculus [1], which refers to the somatosensory and motor body maps in the brain, as those factors are all predetermined. However, using predetermined primitive representations is unacceptable from the viewpoint of neuroscience [60]. Most studies have dealt with discrete-state and discrete-time examples, such as a maze, which is artificial from the neuroscience viewpoint. One rare exception that adopted continuous time and continuous state-space examples was a standing-up task carried out by a multi-joint robot [61], which is categorized as an under-actuated system and considered a difficult system to control as the number of actuators is fewer than the number of degrees of freedom of the robot.

For real-world problems, it is unrealistic to design a hard hierarchy of modules (subgoals) that strictly controls how to switch reward predictions and subgoals between higher and lower hierarchical layers. This limitation suggests that in actual animal reinforcement learning, heterogeneous reward predictions are softly combined, depending on their time-varying significance instead of such a hard hierarchy [62]. The most important aspect of this soft hierarchical architecture is that reward prediction made by limbic and cognitive loops is propagated to motor loops by spiral projections between the striatum and substantia nigra. This mechanism allows reward prediction error, coarsely represented by the cognitive loop, to guide the learning of more detailed reward predictions in motor loops, which incorporate more detailed information, such as motor commands. This soft hierarchical structure shows a useful approach to propagate reward information from the upper to the lower layer of a hierarchical policy-learning mechanism.

In §3, we explore how these modular and hierarchical architectures can be applied to actual humanoid robot control.

## 3. Creating a brain by brain-motivated humanoid-motor-learning methods

An optimal control algorithm for a particular body dynamics may not be optimal for other body dynamics. If a humanoid robot is used for exploring and examining neuroscience



**Figure 3.** Demonstrations of 14 different tasks by humanoid robot DB: imitation learning was involved in ‘Katya-shi’ (Okinawan folk dance) [63]: (a), three-ball juggling [5] (b), devil-sticking (c), the air-hockey demonstration [64,65] (d) uses not only imitation learning but also a reinforcement-learning algorithm with reward (a puck enters the opponent’s goal) and penalty (a puck enters the robot’s goal) and skill learning. Demonstrations of pole-balancing (e) and a visually guided arm reaching towards a target (f) used a supervised learning scheme [66], motivated by our approach to cerebellar internal model learning. Demonstrations of adaptation of vestibulo-ocular reflex [67] (g), adaptation of smooth pursuit eye movement (h) and simultaneous realization of these two kinds of eye movements with saccadic eye movements (i) were based on computational models of eye movements and their learning [68]. Demonstrations of drumming (j), paddling a ball (k), sticky-hand interaction with a human [69] (l), manipulating a box [70] (m) and a tennis swing [71] (n). (Online version in colour.)

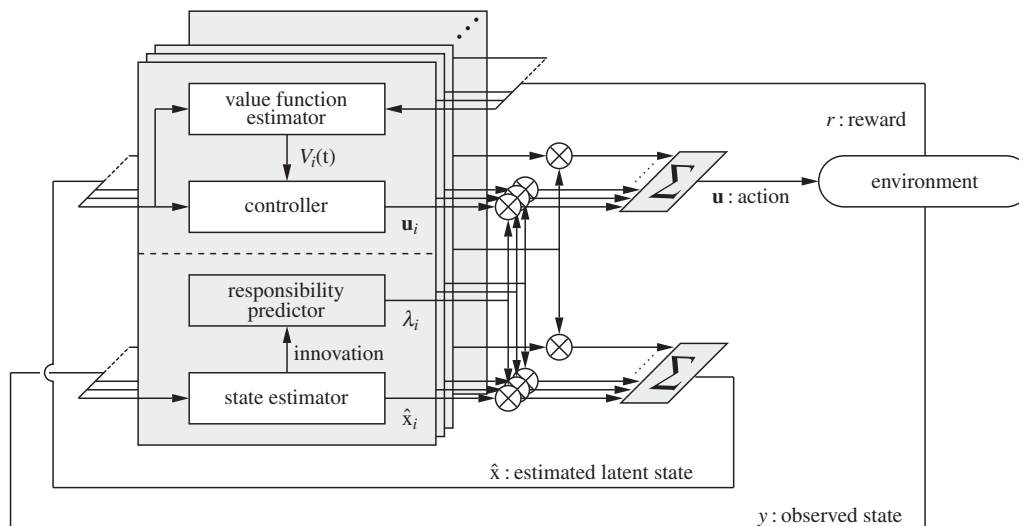
theories rather than for engineering, it should be as close as possible to a human body. Therefore, we first developed a humanoid robot Dynamic Brain (DB) (figure 3), which has 30 d.f. and human-like size and weight. It is equipped with an artificial vestibular organ (gyro sensor), which measures head velocity, and four cameras with vertical and horizontal degrees of freedom. Two of the cameras have telescopic lenses corresponding to foveal vision, and the other two have wide-angle lenses corresponding to peripheral vision. The photographs in figure 3 illustrate 14 of the more than 30 different tasks that can be performed by the DB [5]. Most of the algorithms used for these task demonstrations are approximately based on the principles of information processing in the brain, and many contain some or all of the three learning elements: imitation learning [12,17,72–74], reinforcement learning and supervised learning.

Subsequently, we developed a full-body humanoid robot called CB-i (figure 2a) with the ‘Humanoid Robots as a Tool for Neuroscience’ approach [75]. CB-i was designed for full-body autonomous interaction with balance control and biped walking. Both posture control and biped locomotion are challenging research topics for a human-like and autonomous humanoid robot. Our exploration focused on biologically inspired control algorithms for locomotion using three different humanoid robots (DB-chan [72,76,77], Fujitsu Automation HOAP-2 [48] and CB-i [78]) as well as a small humanoid robot provided by the SONY Corp. [79].

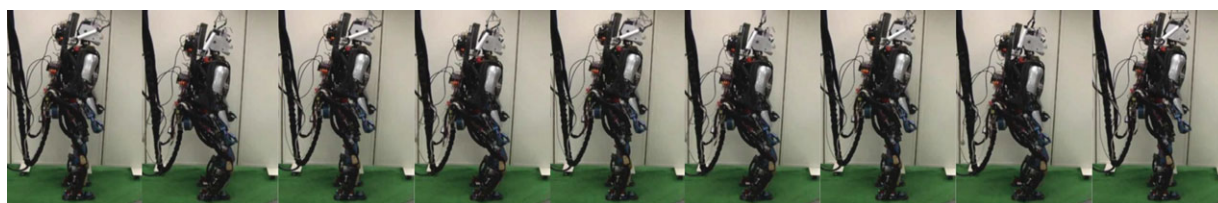
After the above brief introduction of our previous brain-motivated robotics studies, in §3.1, we present our attempts to apply the modular and hierarchical architectures introduced in §§2.1 and 2.2 to humanoid robot platforms.

### 3.1. Modular learning approach for humanoid motor learning

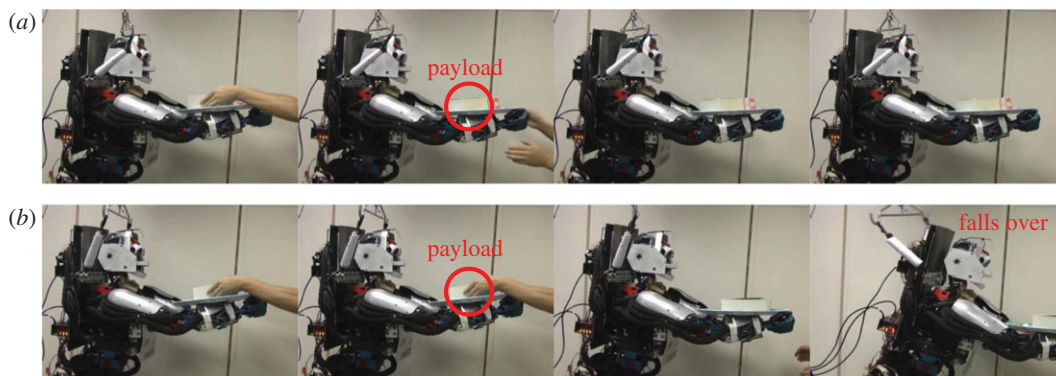
As introduced in §2.1, the MOSAIC architecture is biologically plausible as a human motor-control model. However, two difficulties must be overcome to apply the MOSAIC model to the control of a real humanoid robot: (1) the MOSAIC model does not explicitly consider the existence of noise input to sensory systems and (2) as it assumes full observation, it cannot deal with partially observable systems. To deal with these problems, we proposed an extension of the MOSAIC architecture (figure 4) [80]. Each module of MOSAIC has a forward model, and we can adopt them to construct a state estimator. State estimators can also provide a reasonable model of the sensorimotor function of the central nervous system, as previously suggested [81]. A state estimation strategy using switching linear models is also considered a useful approach for estimating the hidden variables of complicated nonlinear dynamics [82,83]. The modified method, called an extended MOSAIC with state estimators (eMOSAIC), can deal with observation noise and partially observable systems [80]. Based on



**Figure 4.** Extended MOSAIC with state estimators (eMOSAIC) [80]. Each module is composed of a state estimator, a responsibility predictor, a value function estimator and a controller. Nonlinear and non-stationary dynamics are approximated by switching the linear models, and a nonlinear cost function is approximated by switching the quadratic models.



**Figure 5.** Generated humanoid squatting movement using eMOSAIC model [80]. Our humanoid robot maintained its own balance when making a fast squat movement with a frequency up to 1.5 Hz. (Online version in colour.)



**Figure 6.** (a) With the eMOSAIC model, humanoid robot maintained its balance by coping with the sudden environmental change caused by additional payload. (b) With only a single module controller, humanoid robot fell over [80]. (Online version in colour.)

these advantages, our proposed control framework can be applied to real humanoid robots.

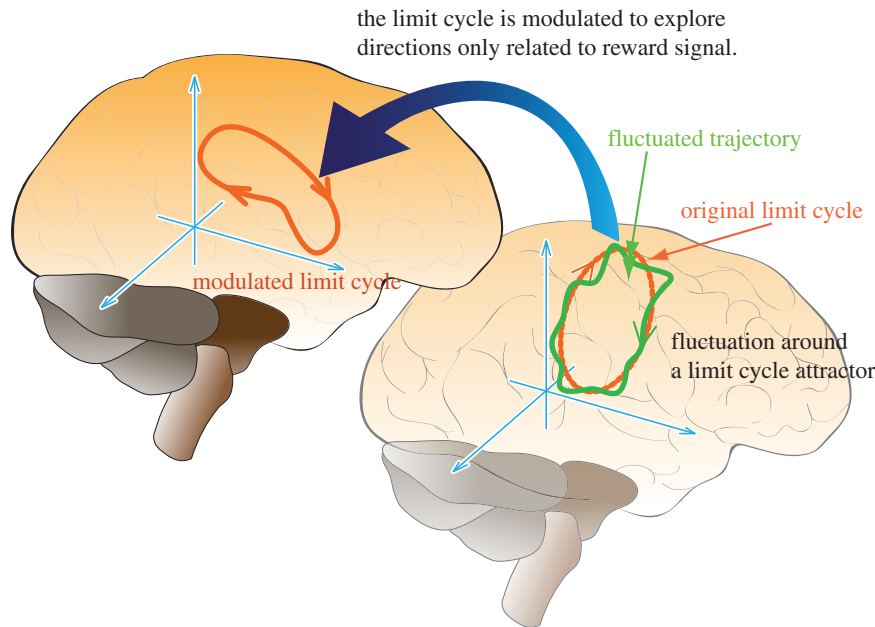
Feedback linearization [84] and gain scheduling control [85] are popular approaches to controlling nonlinear dynamics. However, these methods require either a precise model or prior knowledge of the task and plant structures. On the other hand, for non-stationary systems, adaptive control [86] can be used. However, this method can only handle gradual environmental changes.

We applied eMOSAIC to control our humanoid robot, CB-i (figure 2a), which has highly nonlinear dynamics, to demonstrate that our proposed model can be used in a real environment. Figure 5 shows that a humanoid robot can maintain its balance while performing a squatting task using the

eMOSAIC model. Figure 6 shows that a humanoid robot can maintain its balance even when sudden environmental changes are caused by an additional payload.

### 3.2. Hierarchical architecture for humanoid motor learning

In §2.2, we introduced computational models that use hierarchical architectures to cope with the policy-learning problem in high-dimensional systems. From here, we introduce our idea that the brain modulates the low-dimensional attractor dynamics embedded in high-dimensional state space, where the state space is represented by neural activities (figure 7). The low-dimensional internal dynamics can be task specific



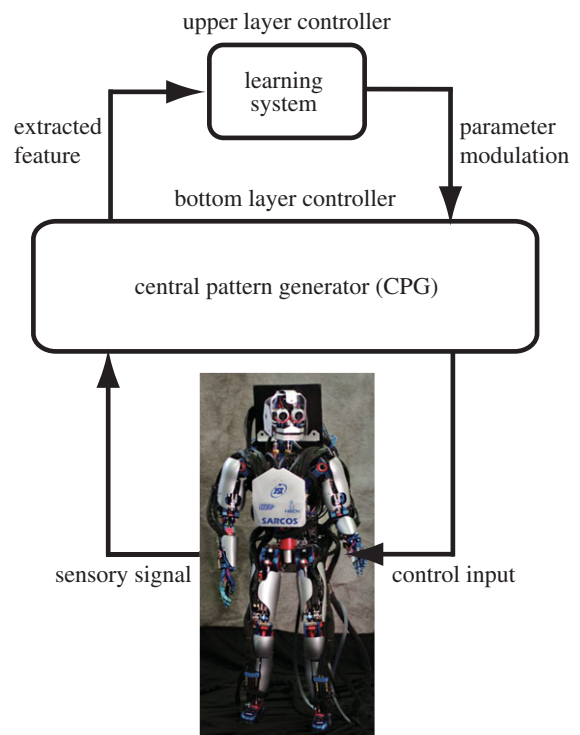
**Figure 7.** Proposed approach to understanding how brain deals with high-dimensional policy-learning problem. We assume low-dimensional limit cycles are represented by super high-dimensional neural activities. Low-dimensional fluctuated trajectories are generated around the limit cycle by noisy inputs. Limit cycle is modulated to explore directions only related to reward signals. (Online version in colour.)

and modulated through the policy-learning process. In this review, as a target task, we generate biped walking movement by humanoid robots with a central pattern generator (CPG) model [87] that is used as the bottom-layer controller to generate the low-dimensional internal dynamics. Figure 8 shows our idea of a hierarchical learning framework. In this section, we introduce a two-layered hierarchical model, where the bottom layer represents the CPG and the upper layer learns to modulate the CPG's periodic pattern to improve the control performance. In §3.2.1, we introduce our CPG model and present feature extraction mechanisms for connecting the upper learning and bottom CPG modules in §3.2.2.

### 3.2.1. Central pattern generator as internal dynamics

As presented in figure 8, we used a CPG model as the bottom-level controller of our proposed hierarchical framework. Here, we introduce our CPG model composed of coupled phase oscillators, where the CPG is believed to exist in vertebrates [88,89]. As it has been emphasized that humanoid robots have inverted pendulum dynamics, where the top is at the centre of mass (COM) and the base is at the centre of pressure (COP) (figure 9), we proposed using the COP to detect the phase of the inverted pendulum dynamics. Using the inverted pendulum approximation to a humanoid model is a popular approach in biped controller design [91,92]. In the proposed biped controller [87], we used simple periodic functions (sinusoids) as desired joint trajectories and showed that the synchronization of the desired trajectories at each joint with inverted pendulum dynamics can generate stepping and walking. In addition, as our nominal gait patterns are sinusoids, our approach does not require careful design of the desired gait trajectories.

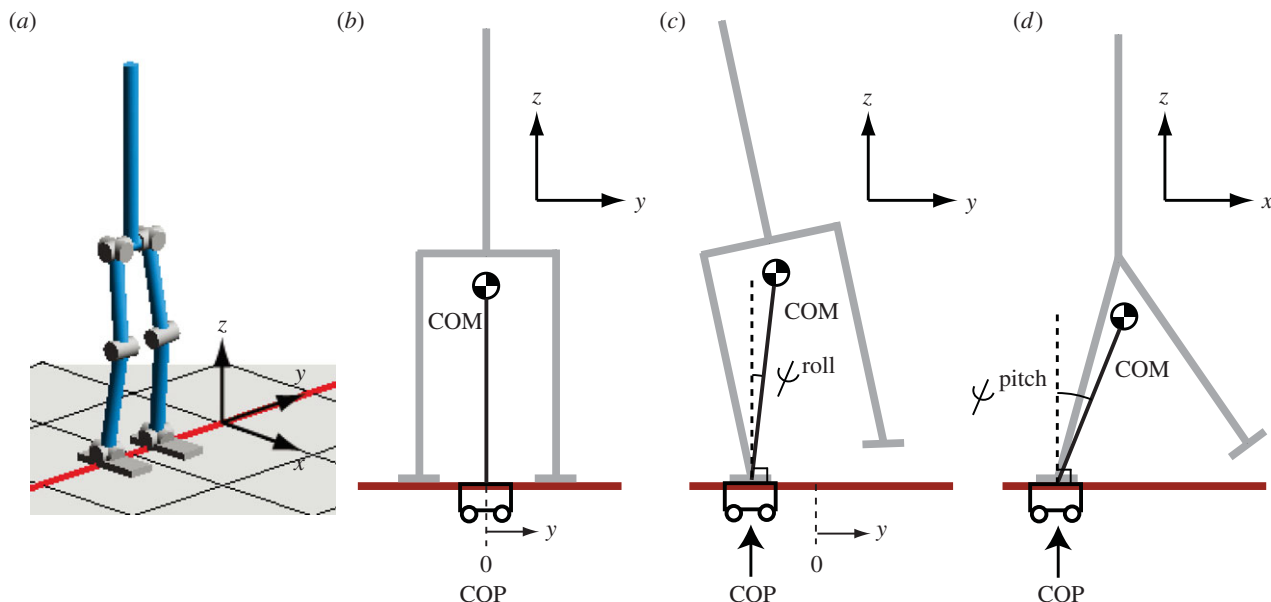
To evaluate the proposed approach in real environments, we applied our proposed biped control method to CB-i. Figure 10 shows its successful walking patterns. The robot walked by using only a simple sinusoidal trajectory, which was composed of at most two sinusoidal basis functions at each joint, modulated by the detected phase from the COP [87].



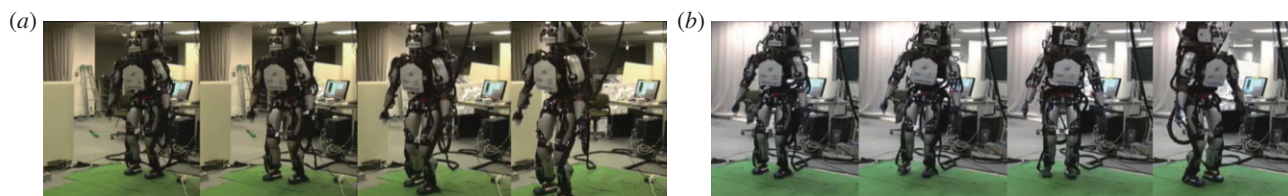
**Figure 8.** Schematic diagram of our biped learning scheme. The CPG generates periodic patterns for biped locomotion. Detected sensory signals are sent to the CPG and the learning system, which only updates the policy and outputs the modulation command to the CPG when the periodic trajectory passes through the defined Poincare sections. (Online version in colour.)

### 3.2.2. Feature extraction for policy improvement

For constructing a hierarchical system, feature extraction mechanisms for connecting the upper learning and bottom CPG modules are necessary. In this section, we introduce our approach to extract low-dimensional feature space for policy updates using kernel dimension reduction (KDR) [93]. As the bottom-level CPG controller can generate the



**Figure 9.** (a) Simplified biped model of CB-i [90]. (b–d) Inverted pendulum model represented by COP and COM.  $\psi^{\text{roll}}$  denotes roll angle of the pendulum.  $\psi^{\text{pitch}}$  denotes its pitch angle. (Online version in colour.)



**Figure 10.** (a) Successful walking pattern of CB-i. (b) Turning movement. (Online version in colour.)

periodic movements, we only consider the robot state at a Poincaré section [94] for policy updates [90].

Low-dimensional feature space needs to contain the information to predict state variable  $\mathbf{y}$ , which represents reward function  $r(\mathbf{y})$ , where the policy parameters are updated to increase the accumulated reward in the reinforcement learning framework [11]. We then formulate the problem to find latent state  $z \in \mathbf{R}^n$  as

$$P(\mathbf{y}_{k+1}|\mathbf{x}_k, \mathbf{u}_k) \simeq P(\mathbf{y}_{k+1}|z_k, \mathbf{u}_k), \quad (3.1)$$

where  $\mathbf{x} \in \mathbf{R}^m$  denotes the state vector in the original state space [90].  $\mathbf{u}$  is the output of a policy. We consider a projection of the state to low-dimensional feature vector  $z = \mathbf{B}\mathbf{x}$  such that  $n < m$ , where  $\mathbf{B}$  is a projection matrix. Then in our feature extraction framework, we consider three sets of state variables: (1) those used to represent the reward function in the original state space:  $\mathbf{y}$ ; (2) the original state variables:  $\mathbf{x}$  and (3) those in the feature space after projection:  $z$ . The relationship in equation (3.1) implies that low-dimensional feature space  $z_k$  maintains the Markov property for state  $\mathbf{y}$ . We used the dimension reduction method, KDR, to derive projection matrix  $\mathbf{B}$ .

We applied our proposed method to improve the walking performance of a simplified model of CB-i, which has 13 d.f., excluding the three translational degrees of freedom of the base link (figure 9a). Therefore, we need to consider a 26-dimensional state space as the original state space that includes the time derivative of each degree of freedom. We applied KDR for the original 26-dimensional state space to find the proper projection

to a one-dimensional feature space. The number of dimensions of the feature space is predetermined.

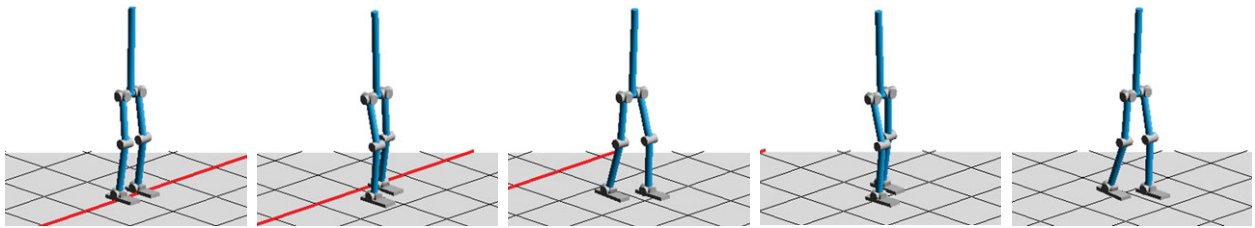
We modulated the amplitude of the hip-pitch sinusoidal joint angles movement to generate forward movement for the biped walking task [90]. The walking task's target is to increase the angular velocity of pendulum  $\dot{\psi}^{\text{pitch}}$  (figure 9d) at section  $\dot{\psi}^{\text{roll}} = 0$  (figure 9c). We used the stochastic Gaussian policy model [95] in which exploratory outputs are generated to sample data for policy improvement. The mean and variance parameters of the Gaussian policy were updated using a policy gradient method [96]. Figure 11 shows the acquired walking performance.

## 4. Interacting with brain through brain–machine interface exoskeleton robots

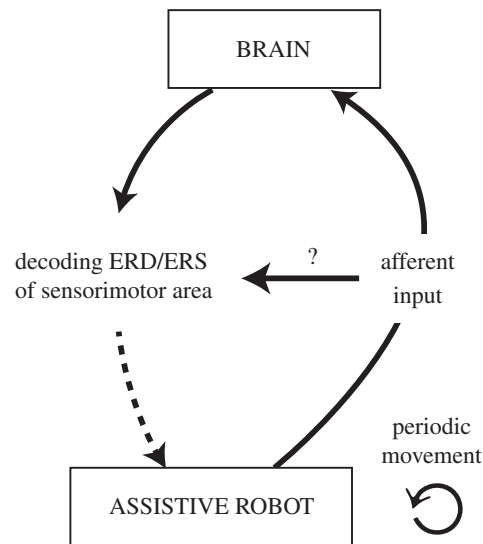
In the previous section, our discussion focused on a possible learning algorithm that modulates a low-dimensional limit cycle embedded in high-dimensional state space. In this section, we turned our focus to introducing our hardware development that modulates user brain activities through physical interaction with our wearable exoskeleton robot.

As many countries are facing the problem of aging populations, the development of an exoskeleton robot to assist user movements is becoming an important research topic [21,97–99]. These exoskeleton robots can be used as assistive or prosthetic devices for stroke and spinal-cord injury patients in rehabilitation programmes [100,101]. In recent years, it has been found that using brain activity to control





**Figure 11.** Acquired walking pattern [90]. Thick (red) line represents starting position. We showed one snap in three walking steps. (Online version in colour.)



**Figure 12.** Investigating how afferent input, induced by an assistive robot, influences the decoding of the ERD/ERS of the sensorimotor area [104]. (Online version in colour.)

a robotic assistive system is also useful to help stroke patients recover their motor functions [33]. This rehabilitation approach is called BMI rehabilitation [102].

For BMI rehabilitation, we have developed an EEG-exoskeleton robot system, where an exoskeleton robot is connected to an EEG measurement device that users can control with their brain activities [44]. To activate an exoskeleton robot based on the user's brain activity, we simultaneously need to consider the influence of the exoskeleton robot movement to the brain activity related to somatosensory inputs as there is physical interaction between the robot and the user. Here, we first introduce our study that investigates whether movement intention can be extracted from EEG signals even when the user movement is assisted by the exoskeleton robot. We then introduce our attempt that used an EEG-exoskeleton robot to assist stand-up movements while it was controlled by the user's brain activity.

#### 4.1. Influence of afferent input induced by an exoskeleton robot

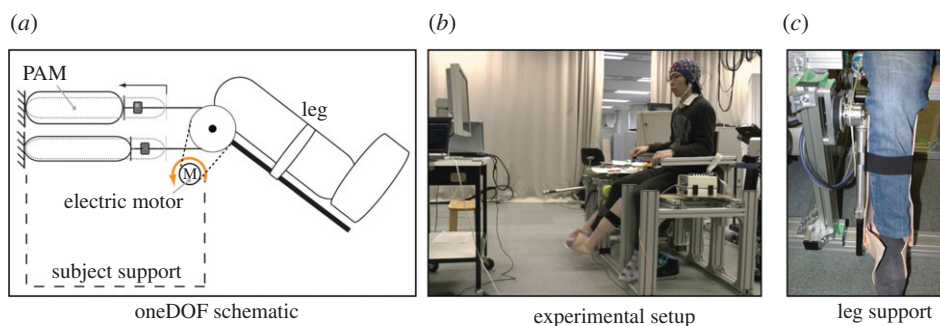
Previous BMI-based robot control studies have dealt with EEG-based neuroprosthesis control [103]. On the other hand, we are focusing on the possibility of interference between the brain activities associated with passive and active movements (figure 12).

In particular, we monitored event-related (de)synchronization (ERD/ERS) to control an EEG-exoskeleton system. The ERD/ERS phenomenon is not only related to active movements or motor imagery but also to passive movements (up arrow in

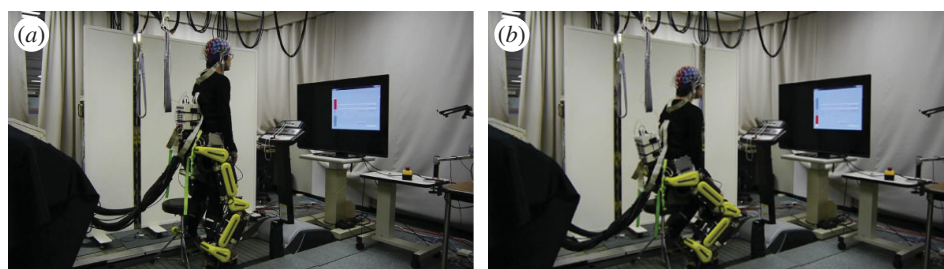
figure 12) that highlight that the feet's passive movements produce a significant ERD/ERS over the entire sensorimotor cortex [105,106]. This raises the question whether the somatosensory afferent input, induced by the periodic leg perturbation, interferes with the decoding ability of a BMI system based on the ERD/ERS of the sensorimotor area (horizontal arrow in figure 12). Specifically, we investigated whether the periodic perturbation of lower limbs produces a significant decrease in the classification performance of the actual left- and right-hand finger movements.

We decoded the left and right sensorimotor hand area due to the reliability of the contralateral ERD/ERS spatial distribution [107] and the fact that this approach is exhaustively discussed in the literature [108,109]. Furthermore, the motivation for using real movements instead of motor imagery is that the latter cannot be mechanically measured or visually assessed by experiment operators. Actually, the patterns of  $\mu$  and  $\beta$  de/synchronization associated with actual movements resemble those of motor imagery, but they just differ in magnitude [110].

In this experiment, our custom-made 1 d.f. exoskeleton robot (figure 13) [104] was used as an assistive robot (figure 13a). The simple exoskeleton robot is actuated using a pneumatic-electric hybrid strategy [43]. The advantage of using a pneumatic actuator is that it can exert a very large torque (maximum of 70 Nm) while generating insignificant electromagnetic noise from the point of view of the EEG system. An electric motor also generates parallel small torque (maximum of 5 Nm) to make fast and precise corrections to the torque generated by the pneumatic actuator. For



**Figure 13.** Robot used during our experiment [104]: (a) PAM and electric actuators exert parallel torques to move the leg (1 d.f. design); (b) Experimental set-up; (c) Thermoplastic polymer leg support anchored to robot and holding up subject's leg. (Online version in colour.)



**Figure 14.** EEG-exoskeleton experiment [44]. Subject controlled EEG-exoskeleton system to follow the stand-up/sit-down directions indicated on the display using motor imagery. Gravity compensation of the lower limb model composed of exoskeleton and human lower limbs was activated when the EEG decoding was 'stand-up.' (a) Stand-up state. (b) Sit-down state. (Online version in colour.)

this specific experiment, the robot was mounted on a custom-made support to allow subjects to sit near to it (figure 13*b*). A leg-shaped thermoplastic polymer was anchored to the robot to secure the legs of subjects (figure 13*c*).

Our study investigates the main effect of the leg afferent input, induced by a lower limb assistive robot, on the decoding performance of the ERD/ERS of the sensorimotor hand area. We experimentally compared a finger-tapping decoding performance under conditions of with and without leg perturbation caused by the exoskeleton robot movements. Our experimental results suggest that the classification performance is always above chance, and we observed no intra-subject significant differences between the with and without-perturbation conditions if we used our proposed decoding procedure [104].

#### 4.2. Electroencephalogram-exoskeleton system

Now we introduce our EEG-exoskeleton system. To decode brain activities from EEG signals, we used a logistic regression method with spectral  $l_1$ -norm regularization, which can lead to good generalization performance [111]. We considered assistance for the stand-up movement, which is one of the most frequently made movements in daily life and a standard task movement in rehabilitation training.

We decoded brain activities using the above classification method based on the logistic regression after pre-processing the measured EEG data. The covariance matrix of the processed data was used as the input variable for the classifier. We updated the covariance matrix of the processed EEG signal with a fixed period. According to the classifier's output, the stand-up or sit-down movements of the exoskeleton robot were generated by the torque control method for the exoskeleton robot [44].

Figure 14 shows the results of the EEG-exoskeleton control. A subject tried to control the EEG-exoskeleton

system using motor imagery to follow the movement directions that appeared on a display, where the directions indicated stand-up/sit-down movements. This neurofeedback training procedure resembles a bio-feedback training method [112] or a BCI training approach [113]. The user successfully controlled the exoskeleton robot using his brain activities.

### 5. Discussion: integrating the three fields by decoded neurofeedback

Finally, in this section, we discuss how the future of computational neuroscience in a broader sense could lead to the integration of all three disciplines, including computational neuroscience, brain-motivated robotics and BMI, based on DecNef methodology.

One of the most critical assumptions in neuroscience and brain science is that the human mind (including consciousness) is caused by the brain's neural activity [114]. Most studies of human learning/memory/cognition have concentrated on examining the correlations between behavioural and neural activity changes rather than establishing cause-and-effect relationships. Even in animal studies, the most frequently used technique is examining the temporal correlation between neural activities and certain hypothetical computational variables proposed by the experimenters. The lack of experimental tools for examining cause-and-effect relationships between brain activity and the mind in systems neuroscience has severely constrained its progress and applicability to such practical problems as robotics or clinical issues [36]. To bridge this gap between major concepts and current technology, we developed a new technique that manipulates neural codes [38] by applying a novel online-feedback method that uses decoded fMRI signals. We call this new technique DecNef. We developed DecNef by spatial

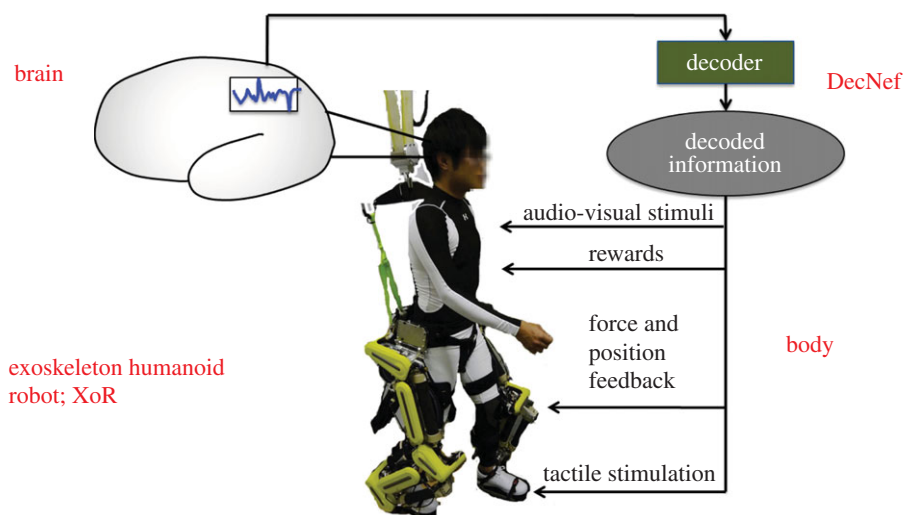
activity patterns in a specific brain area that can be non-invasively controlled by real-time fMRI neurofeedback [38]. We induced visual perceptual learning without visual stimulus presentation or conscious awareness of tasks. The experiment consisted of the following four stages. The first and the fourth are psychological performance tests for the visual discrimination of three orientation gratings. In the second stage, a decoder, actually a machine-learning algorithm called multinomial sparse logistic regression (MSLR) [115], was trained to compute the likelihood of the multi-voxel fMRI patterns that are classified into one of the three orientations used in the above visual discrimination task. Ten years ago, Kamitani & Tong [27] showed that the orientation of seen visual gratings could be 'decoded' from primary visual cortex fMRI multi-voxel patterns. Our DecNef study, which was based on this pioneering study, used MSLR instead of a support vector machine and fMRI signals from the first and secondary visual cortices. The third stage was real-time neurofeedback training that used the above MSLR decoder for orientation to compute visual neurofeedback signals. Participants manipulated their brain activities to obtain maximum monetary rewards. One out of three orientations was allocated to each subject as the target orientation. Six seconds of brain activity were decoded to compute the target orientation's likelihood, which was visually displayed as the size of a green disc 6 s after the brain manipulation. The total amount of monetary reward was proportional to the likelihood, and thus the summation of the disc size across all trials was given to the subject at the conclusion of each day of training. DecNef is essentially a reinforcement-learning paradigm [11] that is guided by monetary reward. It is also related to 'neural operant conditioning.' After 5 or 10 days of DecNef training, when the performances of the first and fourth stages of the experiment were compared, subjects only significantly improved their discrimination performance of the target orientation gratings, but not for the other orientations. Consequently, brain activity as multi-voxel fMRI patterns was experimentally induced without visual stimulus presentation or awareness of the working of the experiment and caused visual perceptual learning. Brain activity was a cause; a behavioural change was the result in this DecNef example.

In a more general framework, DecNef can be described as follows. First, a decoder is constructed that classifies fMRI multi-voxel patterns according to such specific brain information as visual attributes, emotional states or normal and pathological states in psychiatric disorders. In the induction stage, volunteers unconsciously control their brain activity to produce a desirable pattern while being guided by a real-time reward signal computed as decoder output. Thus, the three important elements of fMRI DecNef are the decoding of information from fMRI brain activity, real-time neurofeedback of the decoded information and unconscious reinforcement learning by the participants. DecNef was recently further advanced to manipulate facial preferences [39] or to induce artificial synaesthesia [40]. The basal ganglia was the only brain region that was activated in all three of the above DecNef studies in the induction phase of neurofeedback training; this result is consistent with the reinforcement learning interpretation of DecNef.

Neurofeedback training based on a real-time fMRI has attracted considerable attention for its potential advantages in therapeutic treatment. Most fMRI neurofeedback methods focus on up- or downregulation of a single region-of-interest activation [116–118]. But these have recently been extended

by DecNef to control spatial activity patterns within a specific ROI, as explained above [39]. fMRI–DecNef can be used as a non-invasive BMI system for treating abnormal central decision-making functions in psychiatric disorders [119]. Fukuda *et al.* [120] also demonstrated that fMRI neurofeedback can further modulate spatio-temporal activation patterns across multiple ROIs as well as intrinsic networks. This effect was preserved over a surprisingly long period (more than two months after 3 days' training). Disturbances in intrinsic networks have been reported for a number of neurological and psychiatric diseases. Consequently, the connectivity neurofeedback method could be applied not only as a causal tool for basic neuroscience to alter functional networks but also as a revolutionary therapeutic method by directly curing the pathological connectivity patterns of neurological and psychiatric diseases. DecNef or connectivity neurofeedback [120] might provide revolutionary cures for such psychiatric diseases as autism spectrum disorders for which neither pharmacological nor cognitive/behavioural therapies have proved statistically effective.

DecNef can also be regarded as an extended and specific type of BMI. The neurofeedback itself is one component in a methodology frequently used in BMI. However, DecNef is an extended version of neurofeedback and uses sophisticated decoding and reward feedback, which are not necessarily always involved in BMI. Previous examples of DecNef did not use robots as effectors. BMI directly connects brain activity and an artefact outside the brain for such purposes as sensory substitution, motor compensation and enhancement, neuro-rehabilitation, suppression of neural circuit maladaptation, and cures of brain lesions and neurological and psychiatric disorders. A BMI usually consists of some of the following electrical, mechanical, and computing elements: a brain activity measurement system (electrodes, amplifier, AD converter, etc.), a computer for decoding brain information and/or controlling stimuli to the brain, an effector system (prosthetic limbs, rehabilitation robots, computer cursors, etc.) and a neural stimulation system. DecNef's current form does not depend on the effector or neural stimulation systems, but as outlined below, the future DecNef might also use them to enhance the reward/penalty effects of the DecNef and make training more efficient. In most BMIs, except the sensory substitution BMI, the brain information is first extracted by a decoder from the measured brain signals. This information is transformed and fed back to users in some sensory modality with/without reward signals, and finally BMI users learn to change their brain networks based on synaptic plasticity. Currently, most neurofeedback signals in BMI and DecNef are given in sensory modalities, such as the visual, auditory and somatosensory information of a restricted body part. Reward signals are very important in these neurofeedbacks, especially for DecNef, because DecNef and many BMIs use reinforcement learning as basic working principles. The most compelling neurofeedback device would be an exoskeleton humanoid robot, as it can provide a rich set of sensory as well as kinaesthetic information of all body parts (figure 15). While decoding from a user brain, we can manipulate an XoR to control the spatio-temporal activity patterns of the user's brain to a larger extent than with the currently available narrow neurofeedback channel. One future direction of our work is studying various uses of XoR in neuroscience in addition to BMI rehabilitation. The main target of a combined exoskeleton/DecNef is to experimentally control brain activity patterns either for scientific



**Figure 15.** DecNef and humanoid exoskeleton. Reward signal is crucial in these neurofeedbacks, especially for DecNef. The most compelling neurofeedback device would be an exoskeleton humanoid robot, as it can provide a rich set of sensory as well as kinaesthetic information of all body parts. (Online version in colour.)

purposes or therapeutic interventions. However, inducing plasticity in the brain by DecNef and exoskeleton robots can also improve the control performance of exoskeleton robots by brain activity. We especially need to investigate how the interaction between the human body and the mind with XoR and its control algorithms can produce emotional information that fuel reinforcement learning.

## 6. Conclusion

We introduced our novel ‘creating the brain’ and ‘interacting with the brain’ approaches to understand functional brain mechanisms. In our brain-motivated humanoid motor learning method, we introduced our attempts to understand how the brain deals with high-dimensional reinforcement-learning problems. In our proposed hierarchical learning system, we assumed that low-dimensional limit cycles are represented by

super high-dimensional neural activities. The low-dimensional fluctuated trajectories are generated around the limit cycle by noisy inputs. The limit cycle is modulated to explore the directions that are only related to reward signals. We also introduced a newly developed exoskeleton robot that can be considered a novel robotics tool for neuroscience. As a future direction, we will develop an integrated system of the brain, body and an exoskeleton robot where the robot is controlled by the brain activities to understand how the brain works in natural situations.

**Acknowledgements.** This research was conducted as ‘Development of BMI Technologies for Clinical Application’ under the SRPBS of Japanese MEXT.

**Funding statement.** It was partially supported by ERATO, J.S.T.; by ICORP, J.S.T.; by MIC-SCOPE; by SICIP, J.S.T.; by KAKENHI 2312004; and by contract with the Ministry of Internal Affairs and Communications entitled ‘Novel and innovative R&D making use of brain structures’.

## References

- Kandel E, Schwartz J, Jessell T. 2012 *Principles of neural science*, 5th edn. New York, NY: McGraw-Hill Professional.
- National Institute of Health (NIH). The Brain Initiative. See [www.braininitiative.nih.gov](http://www.braininitiative.nih.gov).
- European Commission. Human Brain Project. See [www.humanbrainproject.eu](http://www.humanbrainproject.eu).
- Marr D. 1982 *Vision*. Cambridge, MA: The MIT Press.
- Atkeson CG *et al.* 2000 Using humanoid robots to study human behavior. *IEEE Intell. Syst. Spec. Issue Human. Robot.* **15**, 46–56.
- Amari S. 1977 Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**, 77–87. (doi:10.1007/BF00337259)
- Doya K. 1999 What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw.* **12**, 961–974. (doi:10.1016/S0893-6080(99)00046-5)
- Fukushima K. 1980 Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.* **36**, 193–202. (doi:10.1007/BF00344251)
- Houk JC, Davis JL, Beiser DG. 1995 *Models of information processing in the basal ganglia*. Cambridge, MA: The MIT Press.
- Kawato M, Gomi H. 1992 A computational model of four regions of the cerebellum based on feedback-error learning. *Biol. Cybern.* **68**, 95–103. (doi:10.1007/BF00201431)
- Sutton RS, Barto AG. 1998 *Reinforcement learning: an introduction*. Cambridge, MA: The MIT Press.
- Schaal S. 1999 Is imitation learning the route to humanoid robots? *Trends Cogn. Sci.* **3**, 233–242. (doi:10.1016/S1364-6613(99)01327-3)
- Kuniyoshi Y, Inaba M, Inoue H. 1994 Learning by watching: extracting reusable task knowledge from visual observation of human performance. *IEEE Trans. Robot. Autom.* **10**, 799–822. (doi:10.1109/70.338535)
- Asada M, Hosoda K, Kuniyoshi Y, Ishiguro H, Inui T, Yoshikawa Y, Ogino M, Yoshida C. 2009 Cognitive developmental robotics: a survey. *IEEE Trans. Autom. Ment. Dev.* **1**, 12–34. (doi:10.1109/TAMD.2009.2021702)
- Rizzolatti G, Craighero L. 2004 The mirror-neuron system. *Annu. Rev. Neurosci.* **27**, 169–192. (doi:10.1146/annurev.neuro.27.070203.144230)
- Calinon S, D’halluin F, Sauser EL, Caldwell DG, Billard A. 2010 Learning and reproduction of gestures by imitation. *IEEE Robot. Autom. Mag.* **17**, 44–54. (doi:10.1109/MRA.2010.936947)
- Miyamoto H, Schaal S, Gandolfo F, Gomi H, Koike Y, Osu R, Nakano E, Wada Y, Kawato M. 1996 A kendama learning robot based on dynamic optimization theory. *Neural Netw.* **9**, 1281–1302. (doi:10.1016/S0893-6080(96)00043-3)
- Ude A, Gams A, Asfour T, Morimoto J. 2010 Task-specific generalization of discrete and periodic dynamic movement primitives. *IEEE Trans. Robot.* **26**, 800–815. (doi:10.1109/TR0.2010.2065430)
- Lungarella M, Matta G, Pfeifer R, Sandini G. 2003 Developmental robotics: a survey. *Connect. Sci.*

- 15, 151–190. (doi:10.1080/09540090310001655110)
20. Birbaumer N, Cohen LG. 2007 Brain–computer interfaces: communication and restoration of movement in paralysis. *J. Physiol.* **579**, 621–636. (doi:10.1113/jphysiol.2006.125633)
21. Dollar AM, Herr H. 2008 Lower extremity exoskeletons and active orthoses: challenges and state-of-the-art. *IEEE Trans. Robot.* **24**, 144–158. (doi:10.1109/TRO.2008.915453)
22. Hochberg LR *et al.* 2006 Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* **442**, 164–171. (doi:10.1038/nature04970)
23. Lebedev MA, Nicolelis MA. 2006 Brain–machine interfaces: past, present and future. *Trends Neurosci.* **29**, 536–546. (doi:10.1016/j.tins.2006.07.004)
24. Nicolelis MA. 2001 Actions from thoughts. *Nature* **409**, 403–407. (doi:10.1038/35053191)
25. Velliste M, Perel S, Spalding MC, Whitford AS, Schwartz AB. 2008 Cortical control of a prosthetic arm for self-feeding. *Nature* **453**, 1098–1101. (doi:10.1038/nature06996)
26. Wolpaw JR, McFarland DJ. 2004 Control of a two-dimensional movement signal by a noninvasive brain–computer interface in humans. *Proc. Natl Acad. Sci. USA* **101**, 17 849–17 854. (doi:10.1073/pnas.0403504101)
27. Kamitani Y, Tong F. 2005 Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* **8**, 679–685. (doi:10.1038/nn1444)
28. Chavarriaga R, Millan JDR. 2010 Learning from EEG error-related potentials in noninvasive brain–computer interfaces. *IEEE Trans. Neural Syst. Rehabil. Eng.* **18**, 381–388. (doi:10.1109/TNSRE.2010.2053387)
29. Cheron G *et al.* 2012 From spinal central pattern generators to cortical network: integrated BCI for walking rehabilitation. *Neural Plast.* **2012**, 1–13. (doi:10.1155/2012/375148)
30. Fetz EE. 1969 Operant conditioning of cortical unit activity. *Science* **163**, 955–958. (doi:10.1126/science.163.3870.955)
31. Hoellinger T *et al.* 2013 Biological oscillations for learning walking coordination: dynamic recurrent neural network functionally models physiological central pattern generator. *Front. Comput. Neurosci.* **2013**, 1–15.
32. Pfurtscheller G, Müller GR, Pfurtscheller J, Gerner HJ, Rupp R. 2003 ‘Thought’–control of functional electrical stimulation to restore hand grasp in a patient with tetraplegia. *Neurosci. Lett.* **351**, 33–36. (doi:10.1016/S0304-3940(03)00947-9)
33. Shindo K, Kawashima K, Ushiba J, Ota N, Ito M, Ota T, Kimura A, Liu M. 2011 Effects of neurofeedback training with an electroencephalogram-based brain–computer interface for hand paralysis in patients with chronic stroke: a preliminary case series study. *J. Rehabil. Med.* **43**, 951–957.
34. Nicolelis MA. Walk Again Project. See [www.releasedigital.com.br/aasdap-iinn-els/walk-again-project](http://www.releasedigital.com.br/aasdap-iinn-els/walk-again-project).
35. MEXT Japan. Strategic Research Program for Brain Science (SRPBS). [brainprogram.mext.go.jp/missionBMI](http://brainprogram.mext.go.jp/missionBMI) (in Japanese).
36. Kawato M. 2008 From ‘understanding the brain by creating the brain’ toward manipulative neuroscience. *Phil. Trans. R. Soc. B* **363**, 2201–2214. (doi:10.1098/rstb.2008.2272)
37. Normile D. 1997 Japanese neuroscience: new institute seen as brain behind big boots in spending. *Science* **275**, 1562–1563. (doi:10.1126/science.275.5306.1562)
38. Shibata K, Watanabe T, Sasaki Y, Kawato M. 2011 Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. *Science* **334**, 1413–1415. (doi:10.1126/science.1212003)
39. Shibata K, Sasaki Y, Kawato M, Watanabe T. 2013 Human facial preferences are changed at the mercy of online neurofeedback. In *Society for Neuroscience Meeting, San Diego, CA, 9–13 November* (Poster presentation).
40. Amano K, Watanabe T, Shibata K, Kawato M, Sasaki Y. 2014 Phenomenal consciousness created by decoded neurofeedback. In *Society for Neuroscience Meeting, Washington, DC, 15–19 November*.
41. Miyamoto H, Kawato M, Setoyama T, Suzuki R. 1988 Feedback-error-learning neural network for trajectory control of a robotic manipulator. *Neural Netw.* **1**, 251–265. (doi:10.1016/0893-6080(88)90030-5)
42. Cheng G, Hyon S, Morimoto J, Ude A, Hale JG, Colvin G, Scroggin W, Jacobsen SC. 2007 CB: a Humanoid research platform for exploring neuroscience. *Adv. Robot.* **21**, 1097–1114. (doi:10.1163/156855307781389356)
43. Hyon S, Morimoto J, Matsubara T, Noda T, Kawato M. 2011 XoR: hybrid drive exoskeleton robot that can balance. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, San Francisco, CA, 25–30 September*, pp. 3975–3981. Piscataway, NJ: IEEE.
44. Noda T, Sugimoto N, Furukawa J, Sato M, Hyon S, Morimoto J. 2012 Brain-controlled exoskeleton robot for BMI rehabilitation. In *Proc. IEEE/RAS Int. Conf. on Humanoid Robots, Osaka, Japan, 29 November to 1 December*, pp. 21–27. Piscataway, NJ: IEEE.
45. Bertsekas DP. 2005 *Dynamic programming and optimal control*. Nashua, NH: Athena Scientific.
46. Haruno M, Wolpert DM, Kawato M. 2001 Mosaic model for sensorimotor learning and control. *Neural Comput.* **13**, 2201–2220. (doi:10.1162/089976601750541778)
47. Wolpert DM, Kawato M. 1998 Multiple paired forward and inverse models for motor control. *Neural Netw.* **11**, 1317–1329. (doi:10.1016/S0893-6080(98)00066-5)
48. Matsubara T, Morimoto J, Nakanishi J, Sato M, Doya K. 2006 Learning CPG-based biped locomotion with a policy gradient method. *Robot. Auton. Syst.* **54**, 911–920. (doi:10.1016/j.robot.2006.05.012)
49. Sugimoto N, Haruno M, Doya K, Kawato M. 2012 Mosaic for multiple-reward environments. *Neural Comput.* **24**, 577–606. (doi:10.1162/NECO\_a\_00246)
50. Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF. 2006 Optimal decision making and the anterior cingulate cortex. *Nat. Neurosci.* **9**, 940–947. (doi:10.1038/nn1724)
51. Hampton AN, Bossaerts P, O’Doherty JP. 2006 The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. *J. Neurosci.* **26**, 8360–8367. (doi:10.1523/JNEUROSCI.1010-06.2006)
52. Haruno M, Kuroda T, Doya K, Toyama K, Kimura M, Samejima K, Imamizu H, Kawato M. 2004 A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *J. Neurosci.* **24**, 1660–1665. (doi:10.1523/JNEUROSCI.3417-03.2004)
53. Samejima K, Ueda Y, Doya K, Kimura M. 2005 Representation of action-specific reward values in the striatum. *Science* **310**, 1337–1340. (doi:10.1126/science.1115270)
54. Haruno M, Kawato M. 2006 Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *J. Neurophysiol.* **95**, 948–959. (doi:10.1152/jn.00382.2005)
55. O’Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. 2004 Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304**, 452–454. (doi:10.1126/science.1094285)
56. Williams ZM. 2006 Eskandar en: selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat. Neurosci.* **9**, 562–568. (doi:10.1038/nn1662)
57. Tanaka S, Samejima K, Okada G, Ueda K, Okamoto Y, Yamawaki S, Doya K. 2006 Brain mechanism of reward prediction under predictable and unpredictable environmental dynamics. *Neural Netw.* **19**, 1233–1241. (doi:10.1016/j.neunet.2006.05.039)
58. Singh S. 1992 Transfer of learning by composing solutions of elemental sequential tasks. *Mach. Learn.* **8**, 323–339. (doi:10.1007/BF00992700)
59. Sutton RS, Precup D, Singh S. 1999 Between MDPs and semi-MDPs: a framework for temporal abstraction in reinforcement learning. *Artif. Intell.* **112**, 181–211. (doi:10.1016/S0004-3702(99)00052-1)
60. Kawato M, Samejima K. 2007 Efficient reinforcement learning: computational theories, neuroscience and robotics. *Curr. Opin. Neurobiol.* **17**, 205–212. (doi:10.1016/j.conb.2007.03.004)
61. Morimoto J, Doya K. 1998 Hierarchical reinforcement learning for motion learning: learning ‘stand-up’ trajectories. *Adv. Robot.* **13**, 267–268. (doi:10.1163/156855399X00513)
62. Haruno M, Kawato M. 2006 Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops; fMRI examination in stimulus-action-reward association learning. *Neural Netw.* **19**, 1242–1254. (doi:10.1016/j.neunet.2006.06.007)
63. Riley M, Ude A, Atkeson CG. 2000 Methods for motion generation and interaction with a humanoid

- robot: case studies of dancing and catching. In *Proc. 2000 Workshop on Interactive Robotics and Entertainment (WIRE-2000)*, Pittsburgh, PA, 30 April to 1 May, pp. 35–42.
64. Bentivegna DC, Atkeson CG, Cheng G. 2004 Learning tasks from observation and practice. *Robot. Auton. Syst.* **47**, 163–169. (doi:10.1016/j.robot.2004.03.010)
65. Bentivegna DC, Atkeson CG, Ude A, Cheng G. 2004 Learning to act from observation and practice. *Int. J. Human. Robot.* **1**, 585–611. (doi:10.1142/S0219843604000307)
66. Schaal S, Atkeson CG. 1998 Constructive incremental learning from only local information. *Neural Comput.* **10**, 2047–2084. (doi:10.1162/089976698300016963)
67. Shibata T, Schaal S. 2001 Biomimetic gaze stabilization based on feedback-error learning with nonparametric regression networks. *Neural Netw.* **14**, 201–216. (doi:10.1016/S0893-6080(00)00084-8)
68. Shibata T, Tabata H, Schaal S, Kawato M. 2005 A model of smooth pursuit based on learning of the target dynamics using only retinal signals. *Neural Netw.* **18**, 213–225. (doi:10.1016/j.neunet.2005.01.001)
69. Hale JG, Pollick FE. 2005 ‘sticky hands’: learning and generalisation for cooperative physical interactions with a humanoid robot. *IEEE Trans. Syst. Man Cybern.* **35**, 512–521. (doi:10.1109/TSMCC.2004.840063)
70. Pollard NS, Hodgins JK, Riley MJ, Atkeson CG. 2002 Adapting human motion for the control of a humanoid robot. In *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA2002)*, Washington, DC, 11–15 May, pp. 1390–1397. Piscataway, NJ: IEEE.
71. Ijspeert AJ, Nakanishi J, Schaal S. 2002 Movement imitation with nonlinear dynamical systems in humanoid robots. In *Proc. IEEE Int. Conf. on Robotics and Automation (ICRA2002)*, Washington, DC, 11–15 May, pp. 1398–1403. Piscataway, NJ: IEEE.
72. Nakanishi J, Morimoto J, Endo G, Cheng G, Schaal S, Kawato M. 2004 Learning from demonstration and adaptation of biped locomotion. *Robot. Auton. Syst.* **47**, 79–91. (doi:10.1016/j.robot.2004.03.003)
73. Ude A, Atkeson CG. 2003 Online tracking and mimicking of human movements by a humanoid robot. *J. Adv. Robot.* **17**, 165–178. (doi:10.1163/156855303321165114)
74. Ude A, Atkeson CG, Riley M. 2004 Programming full-body movements for humanoid robots by observation. *Robot. Auton. Syst.* **47**, 93–108. (doi:10.1016/j.robot.2004.03.004)
75. JST-ICORP. Computational Brain Project, 2004.4–2009.3. See [www.cns.atr.jp/icorp/indexe.html](http://www.cns.atr.jp/icorp/indexe.html).
76. Morimoto J, Atkeson CG. 2003 Minimax differential dynamic programming: an application to robust biped walking. In *Advances in neural information processing systems 15* (eds B Suzanna, T Sebastian, O Klaus), pp. 1563–1570. Cambridge, MA: MIT Press.
77. Morimoto J, Atkeson CG. 2007 Learning biped locomotion: application of Poincaré-map-based reinforcement learning. *IEEE Robot. Autom. Mag.* **14**, 41–51. (doi:10.1109/MRA.2007.380654)
78. Morimoto J, Endo G, Nakanishi J, Hyon S, Cheng G, Bentivegna DC, Atkeson CG. 2006 Modulation of simple sinusoidal patterns by a coupled oscillator model for biped walking. In *IEEE Int. Conf. on Robotics and Automation (ICRA2006) Proc.*, Orlando, FL, 15–19 May, pp. 1579–1584. Piscataway, NJ: IEEE.
79. Endo G, Morimoto J, Matsubara T, Nakanishi J, Cheng G. 2008 Learning CPG-based biped locomotion with a policy gradient method: application to a humanoid robot. *Int. J. Robot. Res.* **27**, 213–228. (doi:10.1177/0278364907084980)
80. Sugimoto N, Morimoto J, Hyon S, Kawato M. 2012 The mosaic model for humanoid robot control. *Neural Netw.* **29**, 8–19. (doi:10.1016/j.neunet.2012.01.002)
81. Wolpert DM, Ghahramani Z, Jordan MI. 1995 An internal model for sensorimotor integration. *Science* **269**, 1880–1882. (doi:10.1126/science.7569931)
82. Ariki Y, Hyon S, Morimoto J. 2013 Extraction of primitive representation from captured human movements and measured ground reaction force to generate physically consistent imitated behaviors. *Neural Netw.* **40**, 32–43. (doi:10.1016/j.neunet.2013.01.002)
83. Ghahramani Z, Hinton GE. 2000 Variational learning for switching state-space models. *Neural Comput.* **12**, 831–864. (doi:10.1162/089976600300015619)
84. Slotine JE, Li W. 1990 *Applied nonlinear control*. Englewood Cliffs, NJ: Prentice Hall.
85. Rugh WJ. 1991 Analytical framework for gain scheduling. *IEEE Control Syst. Mag.* **11**, 79–84. (doi:10.1109/37.103361)
86. Astrom KJ, Wittenmark B. 1989 *Adaptive control*. Boston, MA: Addison Wesley.
87. Morimoto J, Endo G, Nakanishi J, Cheng G. 2008 A biologically inspired biped locomotion strategy for humanoid robots: modulation of sinusoidal patterns by a coupled oscillator model. *IEEE Trans. Robot.* **24**, 185–191. (doi:10.1109/TRO.2008.915457)
88. Grillner S. 1985 Neurobiological bases of rhythmic motor acts in vertebrates. *Science* **228**, 143–149. (doi:10.1126/science.3975635)
89. Ijspeert AJ, Crespi A, Ryczko D, Cabelguen J. 2007 From swimming to walking with a salamander robot driven by a spinal cord model. *Science* **315**, 1416–1420. (doi:10.1126/science.1138353)
90. Morimoto J, Hyon S, Atkeson CG, Cheng G. 2008 Low-dimensional feature extraction for humanoid locomotion using kernel dimension reduction. In *Proc. IEEE Int. Conf. on Robotics and Automation, Pasadena, CA, 19–23 May*, pp. 2711–2716. Piscataway, NJ: IEEE.
91. Kajita S, Hirukawa H, Harada K, Yokoi K. 2014 *Introduction to humanoid robotics*. New York, NY: Springer.
92. Kajita S, Kanehiro F, Kaneko K, Fujiwara K, Yokoi K, Hirukawa H. 2004 Biped walking pattern generation by a simple three-dimensional inverted pendulum model. *Adv. Robot.* **17**, 131–147. (doi:10.1163/156855303321165097)
93. Fukumizu K, Bach FR, Jordan MI. 2004 Dimensionality reduction for supervised learning with reproducing kernel Hilbert spaces. *J. Mach. Learn. Res.* **5**, 73–99.
94. Strogatz SH. 1994 *Nonlinear dynamics and chaos*. Boston, MA: Addison-Wesley Publishing Company.
95. Doya K. 2000 Reinforcement learning in continuous time and space. *Neural Comput.* **12**, 219–245. (doi:10.1162/089976600300015961)
96. Kimura H, Kobayashi S. 1998 An analysis of actor/critic algorithms using eligibility traces: reinforcement learning with imperfect value functions. In *Proc. 15th Int. Conf. on Machine Learning, Madison, WI, 24–27 July*, pp. 284–292. Burlington, MA: Morgan Kaufmann.
97. Kiguchi K, Hayashi Y. 2012 An EMG-Based Control for an upper-limb power-assist exoskeleton robot. *IEEE Trans. Syst. Man Cybern. B* **42**, 1064–1071. (doi:10.1109/TSMCB.2012.2185843)
98. Kobayashi H, Takamitsu A, Hashimoto T. 2009 Muscle suit development and factory application. *Int. J. Autom. Technol.* **3**, 709–715.
99. Yamamoto G, Toyama S. 2009 Development of wearable-agri-robot: mechanism for agricultural work. In *IEEE/RSJ Int. Conf. on Intelligent Robots and System, St. Louis, MO, 11–15 October*, pp. 5801–5806. Piscataway, NJ: IEEE.
100. Kagawa T, Uno Y. 2009 Gait pattern generation for a power-assist device of paraplegic gait. In *The 18th IEEE Int. Symp. on Robot and Human Interactive Communication, Toyama, Japan, 27 September to 2 October*, pp. 633–638. Piscataway, NJ: IEEE.
101. Suzuki K, Kawamoto MG, Hasegarwa H, Sankai Y. 2007 Intension-based walking support for paraplegia patients with robot suit HAL. *Adv. Robot.* **21**, 1441–1469.
102. Kamatani D, Fujiwara T, Ushiba J, Shindo K, Kimura A, Meigen L. 2010 Study for evaluation method of effect of BMI rehabilitation by using near infrared spectroscopy. In *Neuro 2010, Osaka, Japan, 2–4 September*, pp. P2–f16.
103. Müller-Putz GR, Scherer R, Pfurtscheller G, Neuper C, Rupp R. 2009 Non-invasive control of neuroprostheses for the upper extremity: temporal coding of brain patterns. In *Conf. Proc. IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, 2–6 September*, pp. 3353–3356. Piscataway, NJ: IEEE.
104. Lisi G, Noda T, Morimoto J. 2014 Decoding the ERD/ERS: influence of afferent input induced by a leg assistive robot. *Front. Syst. Neurosci.* **8**, 1–12. (doi:10.3389/fnsys.2014.00085)
105. Cassim F, Monaca C, Szurhaj W, Bourriez JL, Defebvre L, Derambure P, Guieu JD. 2001 Does post-movement beta synchronization reflect an idling motor cortex? *Neuroreport* **12**, 3859–3863. (doi:10.1097/00001756-200112040-00051)
106. Wagner J, Solis-Escalante T, Grieshofer P, Neuper C, Müller-Putz G, Scherer R. 2012 Level of participation in robotic-assisted treadmill walking modulates midline sensorimotor EEG rhythms in able-bodied

- subjects. *Neuroimage* **63**, 1203–1211. (doi:10.1016/j.neuroimage.2012.08.019)
107. Wang Y, Hong B, Gao X, Gao S. 2007 Implementation of a brain–computer interface based on three states of motor imagery. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual Int. Conf. of the IEEE, Lyon, France, 23–26 August*, pp. 5059–5062. Piscataway, NJ: IEEE.
108. Blankertz B, Tomioka R, Lemm S, Kawanabe M, Müller K. 2008 Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Proc. Magazine* **25**, 581–607. (doi:10.1109/MSP.2008.4408441)
109. Guger C, Ramoser H, Pfurtscheller G. 2000 Real-time EEG analysis with subject-specific spatial patterns for a brain–computer interface (BCI). *IEEE Trans. Rehabil. Eng.* **8**, 447–456. (doi:10.1109/86.895947)
110. McFarland DJ, Miner LA, Vaughan TM, Wolpaw JR. 2000 Mu and beta rhythm topographies during motor imagery and actual movements. *Brain Topogr.* **12**, 177–186. (doi:10.1023/A:1023437823106)
111. Tomioka R, Müller KR. 2010 A regularized discriminative framework for EEG analysis with application to brain–computer interface. *NeuroImage* **49**, 415–432. (doi:10.1016/j.neuroimage.2009.07.045)
112. Basmajian JV. 1963 Control and training of individual motor units. *Science* **141**, 440–441. (doi:10.1126/science.141.3579.440)
113. Wolpaw JR. 2007 Brain–computer interfaces as new brain output pathways. *J. Physiol.* **579**, 613–619. (doi:10.1113/jphysiol.2006.125948)
114. Gazzaniga MS, Ivry RB, Mangun GR. 2013 *Cognitive neuroscience: the biology of the mind*, 4th edn. London, UK: W. W. Norton and Company.
115. Yamashita O, Sato M, Yoshioka T, Tong F, Kamitani Y. 2008 Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. *NeuroImage* **42**, 1414–1429. (doi:10.1016/j.neuroimage.2008.05.050)
116. deCharms RC, Maeda F, Glover GH, Ludlow D, Pauly JM, Soneji D, Gabrieli DE, Mackey SC. 2005 Control over brain activation and pain learned by using real-time functional MRI. *Proc. Natl Acad. Sci. USA* **102**, 18 626–18 631. (doi:10.1073/pnas.0505210102)
117. Subramanian L, Hindle JV, Johnston S, Roberts MV, Husain M, Goebel R, Linden D. 2011 Real-time functional magnetic resonance imaging neurofeedback for treatment of Parkinson's disease. *J. Neurosci.* **31**, 16 309–16 317. (doi:10.1523/JNEUROSCI.3498-11.2011)
118. Scheinost D, Stoica T, Saksa J, Papademetris X, Constable RT, Pittenger C, Hampson M. 2013 Orbitofrontal cortex neurofeedback produces lasting changes in contamination anxiety and resting-state connectivity. *Transl. Psychiatry* **3**, e250. (doi:10.1038/tp.2013.24)
119. Kawato M. 2013 Computational neuroscience approach to biomarkers and treatments for psychiatric diseases. In *World Congress of Biological Psychiatry, Kyoto, Japan, 23–27 June*. Plenary Lecture, WFSBP.
120. Fukuda M, Kawato M, Imamizu H. 2011 Unconscious operant conditioning of neural activity with real-time fMRI neurofeedback training and its long-term effect on resting state activity. In *Society for Neuroscience Meeting, Washington, DC, 12–16 November*.