Unsupervised Learning of a Reduced Dimensional Controller for a Tendon Driven Robot Platform

Hugo Gravato Marques[1,2], Philip Schaffner[1], and Naveen Kuppuswamy[1]

[1] University of Zurich,
Institute for Informatics, AI Lab.
Zurich 8050, Switzerland
[2] ETH,
Dep. of Mechanical and Process Engineering, BIRL,
Zurich 8092, Switzerland
hgmarques@gmail.com

Abstract. In this paper we present a developmental framework to carry out goal-oriented learning in a low-dimensional space. The framework uses two stages of learning: one to synthesise a set of motor synergies and reduce the dimensionality of the control space in an unsupervised manner, and another to carry out supervised learning in the reduced control space. We test our framework in a reaching task carried out on a (real) tendon-driven robot actuated by four artificial muscles. Our results show that the robot is capable of learning to reach using a reduced control space using no prior information about its body apart from that inherent to the unsupervised and supervised learning rules.

1 Introduction

Current theories of biological motor control in mammals feature almost invariably some kind of hierarchical and modular architecture, where low-level motor primitives (also called stored motor programs, motor units, or muscle synergies), are combined by higher-level mechanisms to produce coordinated behaviour [4][1]. In this context, one of the most prominent areas of research is the identification of motor primitives. This process typically entails some form of factor analysis (e.g. PCA) applied on sensory and motor data collected during behaviour, the outcome of which results on a small number of abstract control units which can explain the behavior of the animal in statistical terms. However, the presence of this statistical units in the animal is difficult to justify or validate, and many open questions remain of whether motor primitives are innate or can emerge from experience [10], [7].

[1] The research leading to these results has received funding from the European Community's 7th Framework Programme FP7 Cognitive Systems, Interaction, Robotics - under grant agreement no. 207212 – eSMCs – and 235065 – RobotDoC, Marie Curie ITN, and the Swiss NCCR Robotics. We thank Cristiano Alessandro for constructing the robot platform used in our experiments.

T. Ziemke, C. Balkenius, and J. Hallam (Eds.): SAB 2012, LNAI 7426, pp. 351–360, 2012.

This question is paramount in artificial systems which make use of hierarchical motor architectures to solve issues of redundancy [11]. In artificial systems motor primitives are typically enforced based on the designer's intuition, and established independently of mechanical considerations, to fulfill a certain functional role (as in variants of the subsumption architecture [2]). But this solution is neither convenient (as primitives are synthesised without following any general principle), nor adaptive (as small unpredictable changes in the system dynamics, or in its environment, might render the primitives useless). In principle, a motor primitive should somehow reflect the regularities and the contingencies of the system to be controlled, rather then being designed independently of the body dynamics [8],[11], which is an argument often used in the field of developmental robotics [1],[5].

In this work we will favour the term "motor synergies" over that of "motor primitives" as the former highlights the co-activation of muscles required to achieve a given task. We present a developmental framework that synthesises a set of motor synergies guided by a self-exploration process. The number of synergies obtained is directly given by the task dimensionality which is lower than the number of actuated muscles. A supervised learning strategy is then used to find the appropriate combination of these synergies and to achieve a desired goal. Our framework has been tested successfully on a (real) tendon-driven pendulum robot actuated by four artificial muscles, where the goal of the robot is to reach to different targets in 2D space.

The reminder of this paper is organised as follows. The second section describes our developmental framework. The third section provides the implementation details of each mechanism in the framework. The fourth section describes the experimental results. The fifth section concludes the paper and provides the outlook of our research.

2 Framework Description

The schematic diagram of the framework proposed in this paper is shown in Figure 1. The framework entails two stages of learning: one carried out by an unsupervised learning process (ULP), and the other by a supervised learning process (SLP). The former synthesises the motor synergies based on Hebbian learning, while the latter uses the synthesised synergies to carry out goal-oriented learning in the resulting reduced control space.

The ULP consists of four interacting mechanisms: a musculoskeletal system (and its environment), a peripheral system, a mechanism capable of triggering spontaneous muscle twitches (SMTs), and a Hebbian-learning mechanism. It works as follows. First, SMTs produce spontaneous and independent contractions in individual muscles. Second, each of these contractions produces forces which are propagated through the musculoskeletal system (as well as through the environment where it is embedded). Third, the changes produced in the musculoskeletal system are captured by the various sensor modalities, which (fourth)

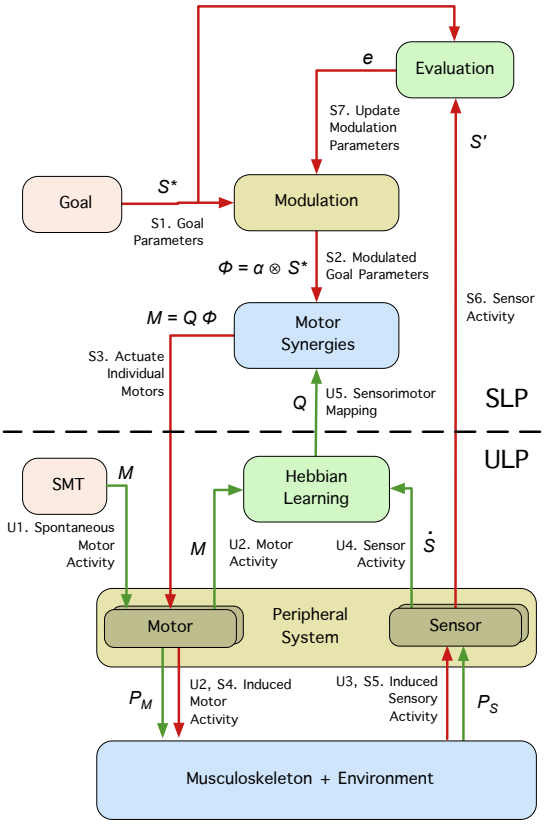
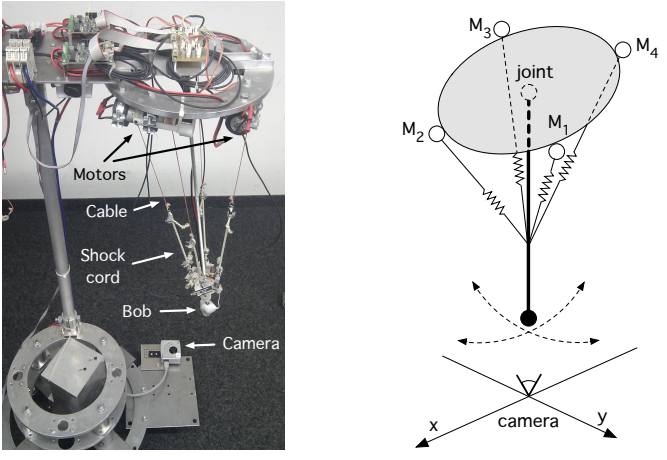


Fig. 1. The proposed developmental framework. The framework entails two processes (divided by a dashed line): a unsupervised learning process (ULP) and a supervised learning process (SLP); the former is shown in gray and the latter is shown in color and it is enclosed by the dashed rectangle. The sequence of events in each process is given by Ui for the ULP and Si for the SLP (see text).

convert them into sensor activity. Fifth, the correlation between the sensor and motor activity is used to synthesise the motor synergies. We have used elsewhere a variant of this scheme to self-organise spinal reflexes [7].

Once the motor synergies are synthesised they are used in the SLP. The SLP consists of an iterative process which aims to identify the appropriate modulation gains required to achieve a given goal. This process entails six mechanisms: a peripheral and a musculoskeletal mechanisms (which are shared with the ULP), a goal mechanism, a mechanism that processes the motor synergies, a modulation mechanism, and an evaluation mechanism. It works as follows. First, a goal is set in the system. Second, the parameters of the goal signals are modulated (i.e. scaled). Third, the modulated signals activate the motor synergies, which



(a) (b)

Fig. 2. The experimental pendulum platform (a) and its schematic diagram (b)

combine the control of the individual muscles. Fourth the actuation of the different motors cause the musculoskeleton to move, which (fifth) induces sensory stimulation (as in the ULP). Sixth, the novel sensor activity is evaluated with respect to the goal. And seventh, the modulation gains are changed and the process is repeated iteratively.

3 Implementation

3.1 Musculoskeletal System and Environment

Our musculoskeletal system consists of a single-joint pendulum actuated by four artificial muscles (see Fig. 2). A camera placed at the bottom tracks the position and velocity of the pendulum. Each muscle consists of a DC Motor a cable and an elastic shock cord arranged in series (see [3],[6]). When the motor is actuated in one direction it reels the cable and creates an analogue to a muscle contraction; when it is actuated in the opposite direction it allows for the muscle to extend. Like their biological counterparts, the artificial muscles have asymmetrical conditioning, i.e. they can only produce force when contracting but not when extending. But unlike biological muscles, our muscles offer a strong resistance to passive extension (i.e. when they are OFF). To overcome this difference we implemented a force controller for each muscle which keeps it (actively) at a minimum tension value when the muscle is supposed to be relaxed.

In both the ULP and the SLP the system starts with all the muscles relaxed, i.e. in minimum tension mode. In this condition the pendulum moves to its resting position due to the effects of gravity. It is noteworthy that the usage of non-standard motors (the motors used are taken from cheap screw-drivers)

combined with the minimum tension procedure introduces variances in the resting position of the pendulum. Nonetheless our architecture is robust to these issues as demonstrated by the results described in Sec. 4.

3.2 Peripheral System

The peripheral system provides the interface between the physical system and the neural (or computational) apparatus; it includes the sensory as well as the motor elements. In our platform, the sensory signal is a 2-element row vector which contains the velocity of the pendulum (obtained from the camera) in the two axes, $\dot{S} = \{\dot{x}, \dot{y}\}$. The derivatives have been filtered using the Savitzky-Golay-filter of 3rd order using a window size of 51. The motor system is 4-element row vector which defines the activity, M_i, of each muscle.

3.3 Unsupervised Learning Process

The single muscle twitches (SMTs) consist of short and spontaneous contractions of single muscles. In mammals this type of motor activity has been observed before birth as well as after birth, during sleep [9]. In our experiments the generation of SMTs is done by sequentially twitching one muscle after the other (each muscle is twitched 20 times). Each twitch consists of a square signal of amplitude $M_i = 4V$ and duration $0.5s$. We have tested different combinations of amplitude and duration and no qualitative impact has been observed on the results presented here.

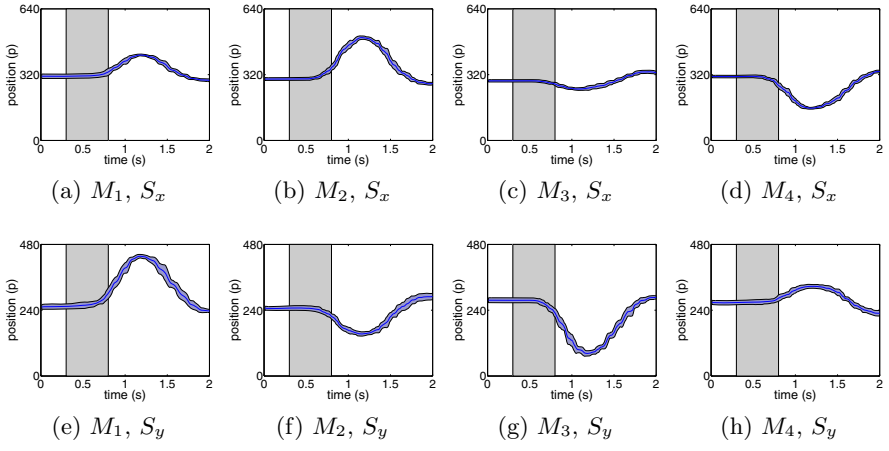
The Hebbian learning mechanism allows to identify the motor synergies based on the correlation between sensor and motor signals during the SMTs . All possible combinations between sensor and motor elements are considered. The motor synergies are described by a single matrix P:

$$P_{i,j} = \eta_{ij} \sum_{t=1}^{T} M_{i,t} \cdot \dot{S}_{j,t+L}, \qquad \eta_{i,j} = \bar{M}_i \left[max(\dot{S}) \sum_{t=1}^{T} M_{i,t} \right]^{-1} \tag{1}$$

where, P is a 4×2 matrix containing information about which motors can be recruited to produce a given sensor activation (positive or negative), η_{ij} is a scaling factor, T is the number of time samples, $M_{i,t}$ is the value of the motor signal M at time sample t, $\bar{M}_i$ is the maximum twitch amplitude of muscle i, $\dot{S}_{j,t}$ is the value of the sensor signal j at time sample t, and L is the lag between sensor and motor activity (obtained *a priori* through direct measurements).

3.4 Supervised Learning Process

In our platform a controller without motor synergies would require a four control signals (one for each motor); the use of muscle synergies reduces the number of controlled variables to only two. This paper investigates a reaching task, where the goal consists of a 2-element row vector, S^* which is defined by $S^* = S_d - S_0$,



(a) M_1, S_x (b) M_2, S_x (c) M_3, S_x (d) M_4, S_x

(e) M_1, S_y (f) M_2, S_y (g) M_3, S_y (h) M_4, S_y

Fig. 3. The raw sensor data collected during the ULP. The data is shown for the x (a-d) and y (e-h) sensor values for SMTs carried out in M_1 (a,e), M_2 (b,f), M_3 (c, g), and M_4 (d,h). Each plot shows the mean and standard deviation of the sensor data for 20 SMTs triggered by the respective muscle. The filled rectangle shows the duration of the twitch.

where S_d is the desired target position of the pendulum, and S_0 is the position of the pendulum at the beginning of each iteration.

The role of the modulation mechanism is to scale the goal signal. Intuitively, a simple multiplication between the goal signal, S^*, and the motor synergies P will move the pendulum in the direction of the goal (see next section). The modulation mechanism allows the amplitude of the movement in the target direction. The modulation mechanism is then given by:

$$\phi = \alpha \otimes S^* \tag{2}$$

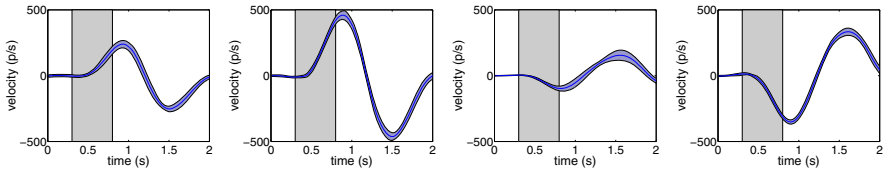
where ϕ is 2-element row vector containing the the modulated goal signals, α is a 2-element row vector containing the modulation gains, and $\otimes$ is the operator for element-wise multiplication.

The motor synergies are responsible to transform the modulated goal signal into individual motor activity. This is done by:

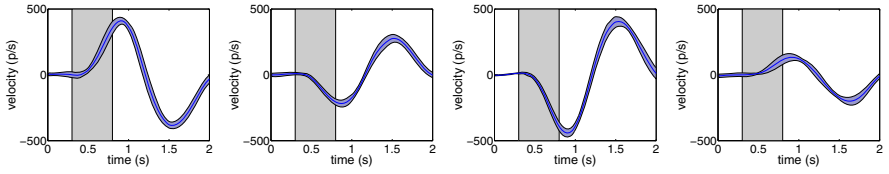
$$M = P\phi \tag{3}$$

The motor signals obtained are activated for a fixed period of time. Here, we used the same time as that used for the twitch duration, i.e. $0.5s$. At the end of each iteration the evaluation mechanism is used to measure the error, e, obtained between the goal vector, S^*, and the movement vector, S'_k, achieved by the pendulum at the end of each iteration. This is given by $S'_k = S_k - S_0$, where S_k is the final position at the end of iteration k. The error in each sensor signal, S, is used to modify the gain parameters according to a gradient descent scheme:

$$\alpha_{k+1} = \alpha_k - \tau \cdot e, \qquad e = (S^* - S') \otimes S^* \oslash |S^*| \tag{4}$$



(a) M_1, $\dot{S}_x$ (b) M_2, $\dot{S}_x$ (c) M_3, $\dot{S}_x$ (d) M_4, $\dot{S}_x$



(e) M_1, $\dot{S}_y$ (f) M_2, $\dot{S}_y$ (g) M_3, $\dot{S}_y$ (h) M_4, $\dot{S}_y$

Fig. 4. The filtered and derived sensor data collected during the ULP. The data is shown for the $\dot{S}_x$ (a-d) and $\dot{S}_y$ (e-h) sensor values for SMTs carried out in M_1 (a,e), M_2 (b,f), M_3 (c, g), and M_4 (d,h). Each plot shows the mean and standard deviation of the filtered and derived sensor data for 10 SMTs triggered by the respective muscle. The filled rectangle shows the duration of the twitch.

where α_k is the gain vector at iteration k, τ is the learning rate (set to 0.15 in our experiments), e is the estimated error in the two dimensions (x and y), $\oslash$ is the operator for element-wise division, and $|S^*|$ is the absolute value of the elements in vector S^*. The error is calculated as the difference between the desired movement vector, S^*, and the vector achieved S'_k; the division between S^* and $|S^*|$ enforces the pendulum to move in the same direction as the target. Note that the only parameters that are being modified during the SLP are the scaling factors, α_k, the dimensionality of which is the same as that of the goal S^* and it is independent of the number of muscles needed to achieve the task.

4 Results

The raw sensor signals S, collected during the ULP are shown in Fig. 3. Each plot shows the mean and the standard deviation of x and y sensor values for a total of 20 SMTs carried out in each muscle. As can be seen SMTs carried out by M_1 increase both the x and y positions of pendulum with respect to the camera. In contrast, SMTs carried out by M_2 increase the x position of the pendulum and decrease the y value. The filtered and derived sensor signals $\dot{S}$ are shown in Fig. 4. As can be observed the profiles match those in Fig. 3; the activation of M_1 produces a positive velocity in both the x and y axes, while the activation of M_2 produces a positive velocity in x and a negative velocity in y. It can also be observed that the relative sensor change is different for each muscle;

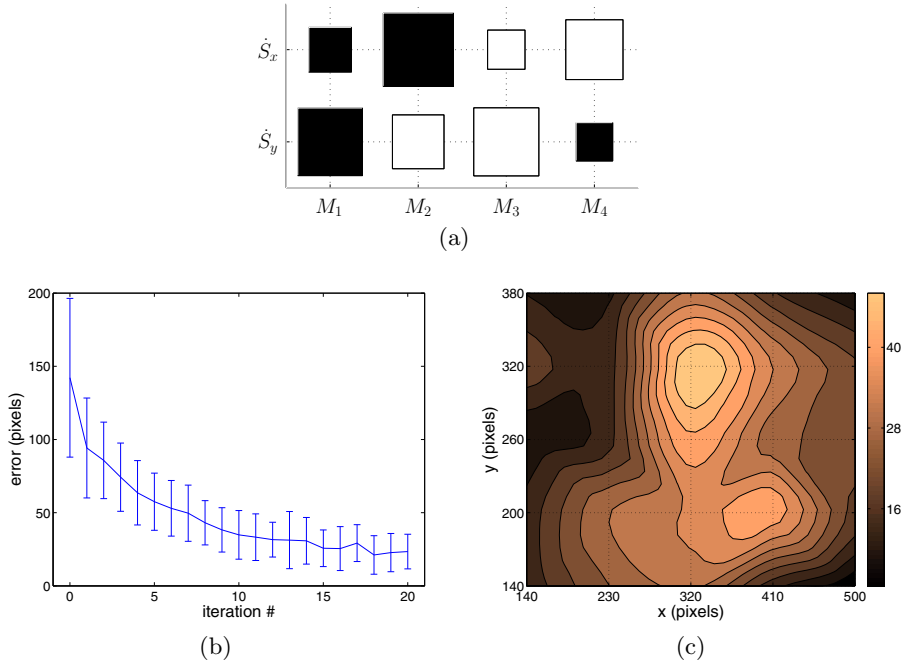
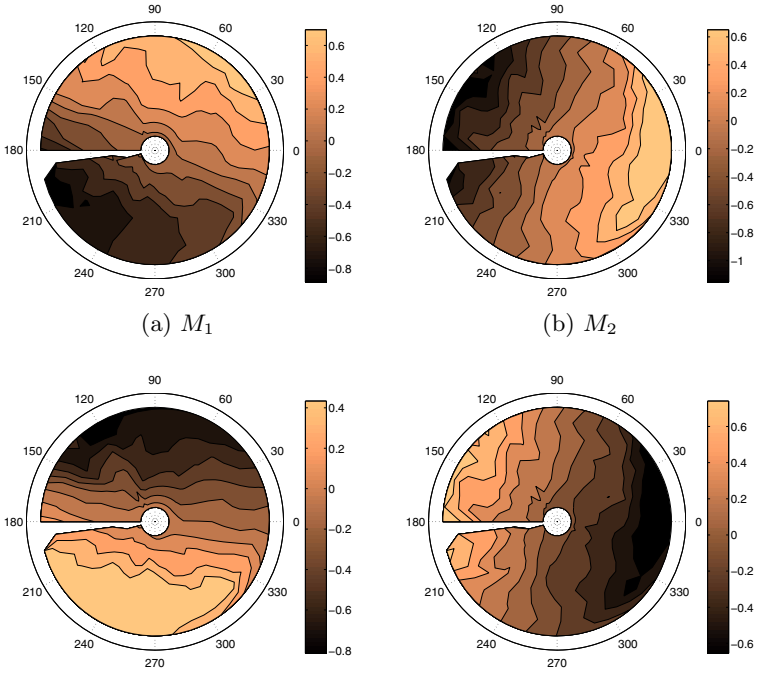


Fig. 5. The motor synergies synthesised during the ULP (a) and the reaching error obtained for the SLP (b-c). a) filled squares represent positive connections and empty squares represent negative connections. The strength of each connection is represented by the size of the respective square, b) the error (mean and standard deviation) obtained for each iteration of the SLP, c) the interpolated error as function of the target position.

for example, SMTs produced by M_1 reach x velocities with smaller amplitudes than those produced by M_2 in the same axis. This is relevant because these amplitudes ultimately affect the weights, $P_{i.j}$ which define the motor synergies.

Matrix P is shown in Fig. 5a (depicted transposed). The resulting connectivity obtained is consistent with the Hebbian learning rule applied to the data in Fig. 4, both qualitatively and quantitatively. From the qualitative point of view, we obtain positive connections between M_1 and both $\dot{S}_x$ and $\dot{S}_y$ signals; this is consistent with the positive velocities reached in both x and y during SMTs produced by M_1. Negative connections can be seen for example between M_2 and $\dot{S}_y$; this is consistent with the negative y velocity achieved during SMTs produced by M_2.

From the quantitative point of view, the connectivity strength is consistent with the signal amplitudes shown in Fig. 4; for example, the connection between M_2 and $\dot{S}_x$ is stronger than that between M_1 and $\dot{S}_x$, which is consistent with the observation that SMTs carried out by M_2 reach higher amplitudes in $\dot{S}_x$ than those carried out by M_1.



(a) M_1 (b) M_2

(c) M_3 (d) M_4

Fig. 6. The motor recruitment as a function of the target position in polar coordinates

Once identified the sensor and motor connectivity using the ULP, we test the SLP by setting a grid of 25 target points and allow the system to make 20 attempts towards each target (each attempt is an iteration in the SLP). The targets are disposed in a grid which tries to maximize the real workspace of the pendulum. At the end of each iteration all the muscles are reset to the minimum tension value to allow the pendulum to move back to its resting position.

Figure 5b shows the mean and the standard deviation of the error achieved when reaching the 25 targets as a function of the iteration step, k. The results show that average of the error decreases rather steadily, although a few error increases can be observed (e.g at iteration 17). These can be explained from 1) the uncertainties of our DC motors, and 2) the fact that the starting position of the pendulum is different at each iteration step (see Section 3.1). The distribution of final error achieved for each target over the entire workspace is shown in Fig. 5c. The smaller errors at the extremities of the workspace can be attributed to the higher precision of the motors when larger voltages are applied.

Figure 6 shows the recruitment of each artificial muscle as a function of the target angle and the target distance. As can be seen the recruitment of each muscle is consistent with its direction of applied force. For example, motor M_1, which pulls in the positive x and y directions (see Figs. 4a,d) is mostly active for

targets located at 45°. In addition, all the motors are mostly active for targets which are farther away, which is in fact expected. Conversely, all the motors are mostly inhibited for targets aligned with the direction of force but located contralaterally to the muscle (e.g. the inhibition of M_1 at 210°).

5 Conclusion

In this paper we have presented a developmental framework to carry out goal-oriented learning in a reduced dimensional space. Our framework was tested in a 2D reaching task carried out by a tendon-driven pendulum robot actuated by four artificial muscles. Our results show that our framework is capable of synthesising a set of motor synergies in an unsupervised manner and combined them effectively to accomplish the proposed reaching task. At the moment we are investigating the use of interpolation methods to generalise the target positions in the task space.

References

1. Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., Origino, M., Yoshida, C.: Cognititive developmental robotics: A survey. IEEE Transactions on Autonomous Mental Development 1(1), 12–34 (2009)
2. Brooks, R.: A robust layered control system for a mobile robot. Technical Report 864, MIT AI Lab (1985)
3. Holland, O., Knight, R.: The anthropomimetic principle. In: Burn, J., Wilson, M. (eds.) Proceedings of the AISB 2006 Symposium on Biologically Inspired Robotics (2006)
4. Latash, M.L.: Evolution of motor control: From reflexes and motor programs to the equilibrium-point hypothesis. Journal of Human Kinetics 19, 3–24 (2008)
5. Lungarella, M., Metta, G., Pfeifer, R., Sandini, G.: Developmental robotics: a survey. Connection Science 15(4), 151–190 (2003)
6. Marques, H., Jäntsch, M., Wittmeier, S., Cristiano, A., Lungarella, M., Knight, R., Holland, O.: Ecce1: the first of a series of anthropomimetic musculoskelal upper torsos. In: Humanoids 2010 (2010) (in press)
7. Marques, H.G., Imtiaz, F., Iida, F., Pfeifer, R.: Self-organisation of reflexive behaviour from spontaneous motor activity (2012) (under review)
8. Pfeifer, R., Lungarella, M., Iida, F.: Self-organization, embodiment, and biologically inspired robotics. Science 318, 1088–1093 (2007)
9. Robinson, S.R., Blumberg, M.S., Lane, M.S., Kreber, L.A.: Spontaneous motor activity in fetal and infant rats is organized into discrete multilimb bouts. Behav. Neurosci. 114(2), 328–336 (2000)
10. Smotherman, W., Robinson, S.: Prenatal ontogeny of sensory responsiveness and learning. In: Greenberg, G., Haraway, M. (eds.) Comparative Psychology: A Handbook, pp. 586–601. Garland Publishing, Inc., New York (1998)
11. Todorov, E., Ghahrammani, Z.: Unsupervised learning of sensory-motor primitives. In: Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2003)