Coding and decoding of information in a bi-directional neural interface

L. Cozzi[1], P. D'Angelo[1], M. Chiappalone[2], A.N. Ide[1,2,3], A. Novellino[2], S. Martinoia[2], V. Sanguineti[1]

[1]*Department of Informatics, Systems and Telematics (DIST), University of Genoa (ITALY)*
[2]*Department of Biophysical and Electronic Engineering (DIBE), University of Genoa (ITALY)*
[3]*Grupo de Processamento de Imagens e Sinais, Universidade Federal de Sao Carlos, Sao Carlos (BRAZIL)*

Abstract

We report on an experiment in which a population of rat cortical neurons, cultured on a micro-electrode array, was connected bi-directionally to a mobile robot. Bi-directional communication between a neural population and an external device requires to translate time-varying signals into spatio-temporal patterns of neural activity, and back. Here we describe the experimental set-up and the computational modules of the neural interface, and work on characterization of the 'transfer function' of the neural preparation, as it emerges from closed-loop experiments, and selection of stimulation and recording sites which are best compatible with a desired behavior.

Keywords: cultured neurons, micro-electrode arrays, neural interface, neural coding

Introduction

Several researchers have proposed to study learning and memory by connecting nervous tissue, kept alive *in-vitro*, bi-directionally to an external device, i.e. an actual physical body. For instance, Reger & al. [1] connected a lamprey brain, isolated and kept alive *in-vitro*, bi-directionally to a mobile robot. The robot was presented light stimuli, and the brain reacted to them by inducing the robot to follow or escape the light source. Similar experiments were performed by DeMarse & al. [2], who interfaced a neuronal network cultured on a micro-electrode array to a computer-simulated animal, moving inside a virtual world. Following this approach, and with the aim of establishing general interfacing techniques and computational

methods [3], we interfaced a mobile robot with a population of neurons, extracted from rat embryos and cultured on a micro-electrode array.

In-vitro cultured neurons form a bi-dimensional physical model of the brain and, in spite of their simplified level of organization, are an useful framework to study information processing in the nervous system. One peculiar feature of this preparation is the possibility to stimulate and/or record from multiple sites at the same time. To establish a neural interface between cultured neurons and an external device, at least three issues need to be addressed: (i) how to optimally translate time-varying, multi-dimensional sensory information into spatio-temporal stimulation patterns; this is a kind of reverse of the neural coding problem - synthesis rather than analysis; (ii) how to efficiently translate spike trains, recorded from multiple neurons, into time-varying 'control commands' for the external device; and (iii) how to induce selective changes in the input-output behavior of the neural preparation; i.e. sensorimotor learning. Embodiment places constraints on the latter, in the sense that puts learning in relation to behavior; i.e., learning has to be aimed at the emergence of a specific behavior. Learning also implies that, to preserve efficiency (i.e., optimality) of the neural 'interface', the processes of coding and decoding should adapt as well. Here we describe the experimental set-up and the computational modules of the neural interface; we show the results of our initial closed-loop experiments, and describe a procedure for identifying the stimulation and recording sites which best predict a desired behavior (obstacle avoidance).

Materials and Methods

Experimental set-up. Primary cultures of cortical neurons, extracted from rat embryos (17-18 days), were plated on planar arrays of 60 TiN/SiN electrodes, arranged along a square grid (MEA60, Multichannel Systems, Reutlingen, Germany), and equipped with an integrated 60-channel pre-amplifier. The diameter of the electrodes and the inter-electrode spacing were, respectively, 30 μm and 200 μm. Experiments started after 18-34 days in-vitro (DIV), i.e.

when these preparations reportedly display maximum spontaneous activity [3]. Stimulation consisted of monopolar, biphasic stimuli (peak-to-peak amplitude 1.5 V, duration 250 μs), generated by a custom stimulator under computer control. These stimulus parameters are optimized in terms of magnitude, reliability of evoked activity, and minimal damage to the preparation [4]. The external body consists of a miniature mobile robot (Khepera II, K-team, Préverenges, Switzerland), equipped with two wheels and eight infrared proximity sensors. For each preparation, we identified two sets of channels to be, respectively, the recording and stimulation sites, i.e. the 'motor' and 'sensory' areas of the model brain. The sensory area simply consisted of two channels, which coded the average activity of the proximity sensors on, respectively, the left and the right side of the robot. Two separate sets of recording sites, eight sites each, were selected to control the left and right wheels of the robot. As the target behavior, we focused on a simple obstacle avoidance task, i.e. a 'Braitenberg vehicle' [5].

Computational architecture. A block diagram of the computational architecture is depicted in Figure 1A.

Figure 1 near here

The time-varying signal which comes from the robot proximity sensors, $u(t)$, is sampled at 10 Hz and averaged into 'left' and 'right' sensor activities, $s_L(t)$ and $s_R(t)$ – this corresponds to having defined two receptive fields – and is translated into, respectively, a left and right instantaneous rate of stimulation, $r_s(t)$. For these experiments, we used a proportional coding scheme, i.e. the rate of stimulation is proportional to sensor activity. The maximum rate of stimulation, r_s^{max}, is only attained when the robot hits an obstacle; this value should be as large as possible for accurate coding of the temporal structure of the sensory signal [4], but there is an upper limit on the energy per unit time that can be delivered to the preparation without damage. In these experiments, we conservatively used r_s^{max} = 2-5 Hz. The instantaneous rate of stimulation directly modulated a perfect integrate-and-fire module. This allows a suffciently reliable coding accuracy, and a coding fraction of about 90%, see [6].

This means that the amount of Left and Right sensor activity that is lost while coding into trains of stimuli is about 10%.

As regards the decoding part of the interface, the raw signals recorded from the sixteen 'Output' sites are sampled at 10 kHz, and individual spikes are detected on-line through a simple peak-to-peak thresholding algorithm, over a 4 ms sliding window. Stimulus artifact is suppressed on each recorded signal by blanking a 8 ms window following each stimulus on either input site.

The instantaneous firing rates, $r^y(t)$ are then estimated on-line from the recorded spike trains, $y(t)$, through a 2nd order low-pass filter, with $f_{cutoff} = 0.1$ Hz. Finally, these resulting spatio-temporal patterns of neural activity have to be translated into a pair of motor commands for the two wheels of the robot. In these experiments, we simply chose to take the average neural activity of the Left and Right portions of the Output region:

$$\omega(t) = \omega_o \left(1 - \frac{1}{2} \frac{1}{N_{out}} \sum_{i=1}^{N_{out}} z_i(t) \right) \qquad (1)$$

where $z(t)$ is the normalized activity:

$$z(t) = C \cdot \hat{r}^y(t) \qquad (2)$$

and C is a diagonal matrix of coefficients. The diagonal terms of C correspond to the reciprocals of the average firing rates evaluated during a preliminary characterization phase:

$$c_{jj} = \frac{1}{\left\langle \hat{r}_j^{\,y} \right\rangle} \qquad (3)$$

The purpose of normalization is to allow, on average, all the recording sites to contribute equally to the motor command, irrespective of their average firing rate. According to the control law of Eq. 1, increased neural activity decreases the speed of the wheels. The maximum angular speed was set to $\omega_o = 4$ rad/s, and the number of recording channels for each portion of the Output region was set to $N_{out} = 8$. The resulting speed commands were finally sampled at 10 Hz and delivered to the robot.

Real-time control. The above described computational architecture was mapped into two P4 PCs. A first PC is responsible for the most computation-intensive portion of the calculations, i.e. acquisition of raw neural data (at a 10 kHz sampling rate) and on-line spike detection. A second PC is responsible for spike decoding, communication with the robot and coding of sensor activity into patterns of stimulation. This part of the architecture runs at 250 Hz (communication with the robot runs at 10 Hz). An additional PC is used as experiment front-end, and a fourth PC, outside the control loop, is used for monitoring and logging the raw recorded neuronal data.

Selection of Input sites. Series of 50 stimuli (0.2 Hz) were delivered to the neural preparation from eight sites (candidate inputs), chosen among those that were most spontaneously active. For each of them, we assessed the average number of evoked spikes. We then discarded the sites that elicited either smallest or largest activity, and selected the pair which elicited the "intermediate" response.

I/O connectivity. To characterize the functional I/O connectivity of the neural preparation, we performed a linear regression of the estimated firing rates at the Output sites, over the sensory signals that are delivered to the neural preparation by stimulating the Input sites.

$$\tilde{r}_y(t+\Delta) \approx H \cdot s(t) \qquad (4)$$

To account for the time lag between stimulus and response, the firing rates has to be shifted of $\Delta \approx 1.5$ s, which is the time lag introduced by low-pass filter used for firing rate estimation. To estimate the regression coefficients, the preparation was stimulated from the two selected inputs, by using as $s(t)$ two random processes (random binary processes, plus Gaussian white noise, to simulate switching between obstacles and free space). The maximum stimulation rate was set to 2 Hz. The resulting coefficients are then normalized with respect to the average firing rate of each channel.

Selection of Output sites. The candidate outputs are the 32 maximally responsive recording sites. Among the 32 candidate outputs, 8 left outputs and 8 right outputs were chosen

according to the functional interconnectivity map in order to result in a behavior that is as close as possible to obstacle avoidance. Obstacle avoidance can be achieved if activation of sensors on one side elicits a decrease of speed at the opposite side. Because of Eq.1, this requires that an increase of stimulation rate on the left (right) stimulation site causes an increase of neural activity on the left (right) group of recording site. To select the recording sites that are most compatible with this behavior, we considered the normalized functional connectivity map:

$$V = C \cdot H \qquad (5)$$

Each row of V corresponds to a recording site, and the two coefficients indicate its response to either left (v_1) or right (v_2) unit stimulation. As a selection criterion, we computed the absolute value of the difference between v_1 and v_2. The best candidates are the recording electrodes that correspond to a higher value of this parameter, i.e. the ones whose position on the v_1- v_2 diagram (see Figure 3A for an example) are most distant from the diagonal, which corresponds to equal response to left and right stimulation.

Closed-loop experiments. The robot was placed inside a circular arena (0.8 m diameter) containing a number of cylinder-shaped obstacles. A total of 4 neural preparations were used in these experiments. Each experiment consisted of four closed-loop trials, each lasting 10 min, separated each other by rest periods.

Results

Figure 1B-G shows a 10-s portion of a typical experiment. The sensory signal tends to switch between two extreme values (no obstacles, or obstacle hit), reflecting the short range of action (about 5 cm) of IR proximity sensors (Figure 1E).

Sensitivity to stimulation. To derive a more synthetic description of the input-output behavior of the neural preparation, as it emerges from these experiments, for each input-output pair of channels we estimated the impulse response of the neural preparation by using a frequency-

domain technique in order to account for auto-correlation of the train of input stimuli. The Fourier transform of the impulse response is computed, in the frequency domain, as the ratio between the cross-spectrum between the input and the output and the power spectrum of the input [6]:

$$H(\omega) = \frac{S_{xy}(\omega)}{S_{xx}(\omega)} \qquad (6)$$

Representative results for one of our experiments are displayed in Figure 2.

Figure 2 near here

This generalizes the post-stimulus time histogram (PSTH), i.e. the average output firing rate in response to an individual stimulus on one of the input channels.

A typical feature of the observed impulse responses is a peak at about 100 ms following the stimulus; the latency of the peak varies little at the different recording sites.

Stability of the response. To test stability of the input-output response of the neural preparation during an individual experimental trial, we repeated the computation of the impulse response by using the first half and the second half of the trial. We found that the impulse responses change little with time across the same trial; see Figure 2.

Functional connectivity. Figure 3A shows the interconnectivity maps obtained in one closed-loop experiment and in the ideal case, corresponding to a Braitenberg 'fear' behavior (inset). In the ideal situation, left and right output sites are arranged into two clearly separated clusters. In contrast, in experimental conditions (number of channel is enclosed in a solid or a dashed box, respectively), the separation between the response of left and right output sites is less evident. This implies a more random connectivity between input and output sites.

The law that converts the sensory inputs s into the motor commands ω, can be expressed as:

$$\omega = \omega_0 (1 - A s) \qquad (7)$$

The matrix A represents the robot controller implemented by the neural preparation:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad (8)$$

The values of its elements are computed as the sum of the coefficients of the normalized connectivity matrix, corresponding to each input/outputs combination (left/left, left/right, right/left, right/right). Obstacle avoidance is achieved if $a_{11}<a_{21}$ and $a_{12}>a_{22}$. Figure 3C shows the diagram of the coefficients of the matrix A that is associated to the output sites selected in the same experiment as in Figure 3A. Their values are compatible with an obstacle avoidance behavior.

Optimal linear reconstruction. The connectivity matrix can be also used to estimate the input signals from the observed instantaneous firing rates. The maximum-likelihood linear estimator of $s(t)$ given H and $r^y(t)$ is defined as:

$$\hat{s}(t) = (H^T \cdot \Sigma^{-1} \cdot H)^{-1} \cdot H^T \cdot \Sigma^{-1} \cdot \tilde{r}_y(t+\Delta) \qquad (9)$$

where Σ is the covariance matrix of the fitting error of the linear model of Eq. 4. Figure 4 (left) shows the comparison between an actual sensory signal and its reconstruction. Figure 4 (right) demonstrates that the reconstruction performance is maximum for a cut-off frequency of 0.1 Hz. This is the reason of the choice of the cutoff frequency of the filter for the estimation of the instantaneous firing rates.

Discussion

We have described a computational architecture, and its real-time implementation, for interfacing a population of cultured neurons to an external device (a mobile robot). We suggested that the PSTH, estimated for each individual input/output pair, may provide useful information on the input/output behavior of the preparation. PSTH is relatively stable with time, and may be an useful tool to assess changes in the input-output behavior induced by adaptation protocols. Furthermore, a 'functional I/O' connectivity map can be defined by calculating the coefficients of the linear regression of the estimated firing rates at the output sites, over the sensory signals coming from the proximity sensors of the robot. These

indicators can be used to identify the input/output channels of the neural preparation that correspond to a motor behavior that is as close as possible to obstacle avoidance.

Acknowledgements

This work is partly supported by the EU grant #IST-2001-33564 (NeuroBIT).

References

[1] Reger B.D., Fleming K.M., Sanguineti V., Alford S., and Mussa-Ivaldi F.A. (2000) Connecting Brains to Robots: An Artificial Body for Studying the Computational Properties of Neural Tissues. *Artificial Life*, 6(4): 307–324.

[2] DeMarse T.B., Wagenaar D.A., Blau A.W. and Potter S.M. (2001) The Neurally Controlled Animat: Biological Brains Acting with Simulated Bodies. *Autonomous Robots*, 11, 305-310.

[3] Martinoia S., Sanguineti V., Cozzi L., Berdondini L., van Pelt J., Tomas J., Le Masson G., Davide F. (2004) Towards an embodied *in-vitro* electrophysiology: the NeuroBIT project. *Neurocomputing*, 58–60: 1065–1072.

[4] Wagenaar DA, Pine J, Potter SM (2004) Effective parameters for stimulation of dissociated cultures using multi-electrode arrays. *J Neurosci Methods* 138(1-2):27-37.

[5] Braitenberg V (1984) Vehicles – Experiments in Syhthetic Psychology. MIT Press.

[6] Gabbiani F, Koch C (1996) Coding of Time-Varying Signals in Spike Trains of Integrate-and-Fire Neurons. *Neural Computation* 8(1): 44-66.

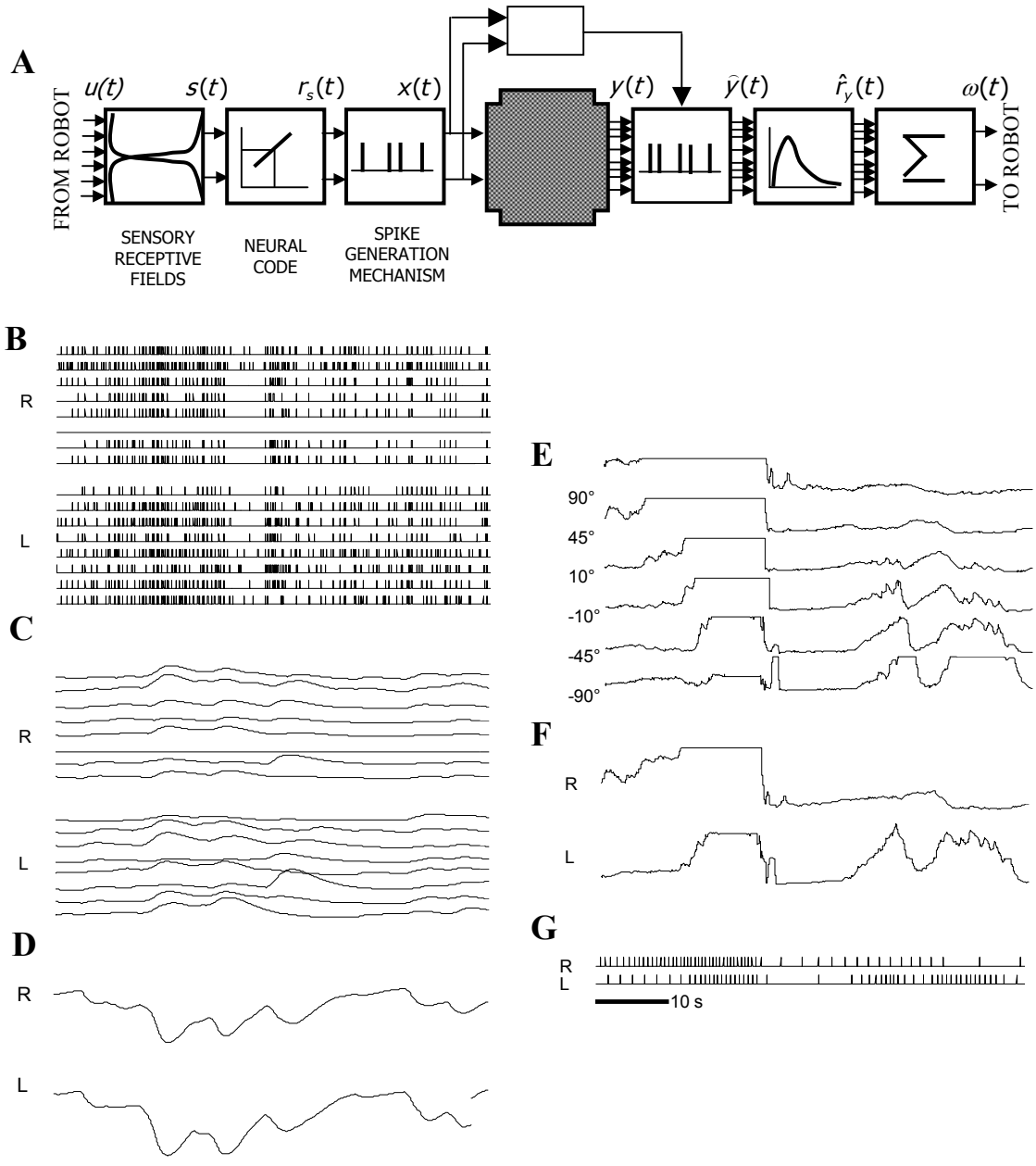
A
u(t)
s(t)
r_s(t)
x(t)
y(t)
ŷ(t)
r̂_y(t)
ω(t)
FROM ROBOT
TO ROBOT
SENSORY RECEPTIVE FIELDS
NEURAL CODE
SPIKE GENERATION MECHANISM
B
R
L
C
R
L
D
R
L
E
90°
45°
10°
-10°
-45°
-90°
F
R
L
G
R
L
10 s

Figure 1

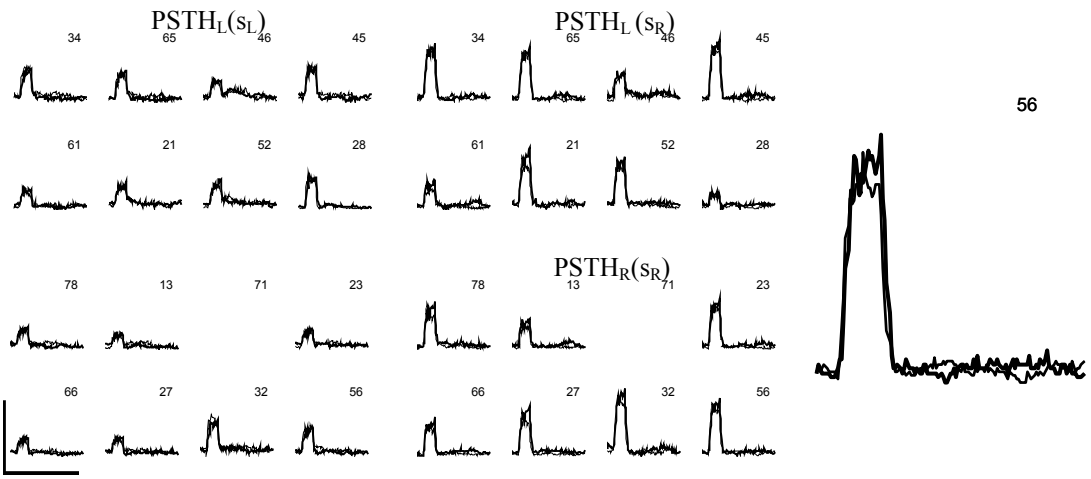
PSTH$_L$(s$_L$)
PSTH$_L$ (s$_R$)
PSTH$_R$(s$_R$)
34
65
46
45
61
21
52
28
78
13
71
23
66
27
32
56
56

Figure 2

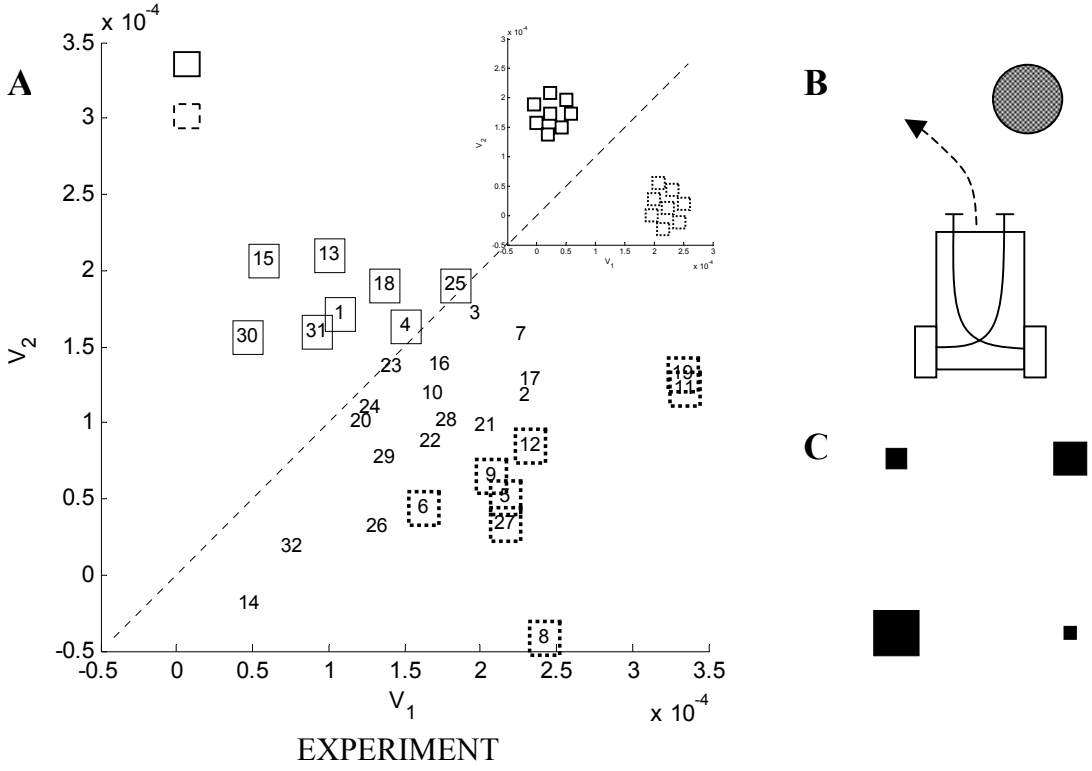

Figure 3

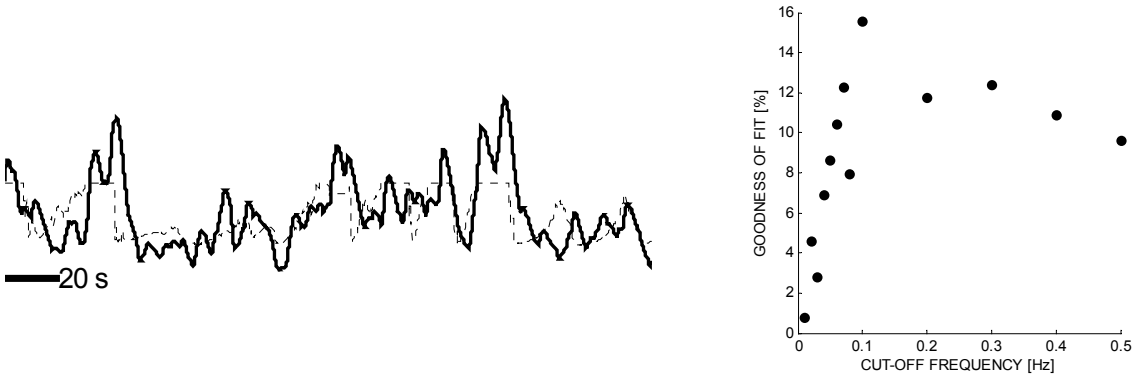
20 s
GOODNESS OF FIT [%]
CUT-OFF FREQUENCY [Hz]
0
2
4
6
8
10
12
14
16
0
0.1
0.2
0.3
0.4
0.5

Figure 4

Figure 1 A: Block diagram of the bi-directional neural interface. B-G: A 40-s portion of a typical closed-loop experiment: recorded spike trains (B) and instantaneous firing rates (C) from the 16 output sites, motor commands (D) and, on the right, the corresponding activity of the IR sensors (E), the average Left and Right sensor activity (F) and the corresponding pattern of stimulation (G).

Figure 2 Impulse responses of the neural preparation for each input-output pairs (bin size: 4 ms)**.** Left, from top to bottom: Responses to stimulation of Left channel. Center, from top to bottom: Responses to stimulation of Right channel. Each impulse response is estimated from the first half (thin lines) and the second half of the data (thick lines). Channel 71 does not respond to stimulations on either side. Right: magnification of the impulse response of channel 56 (R output region, stimulation of R side). Scale bars: 0.5 s, 100 spikes/s.

Figure 3 Coefficients of the functional interconnectivity map of left (dashed) and right (solid) output sites in response to stimulation of left vs. right input channels. A. Results of the clustering algorithm applied to our experimental data. Insert: Ideal clustering of the coefficients, for achieving obstacle avoidance. B: Braitenberg 'fear' behavior, leading to obstacle avoidance. C. Diagram of the coefficients of the controller corresponding to the functional interconnectivity map in A: the size of each square is proportional to the value of the coefficient.

Figure 4 Left: Comparison between an actual sensory signal and its reconstruction obtained by means of the maximum-likelihood linear estimator. Right: Performance in reconstruction as a function of the cutoff frequency used for estimating the instantaneous firing rate.