

# Rewards-Driven Control of Robot Arm by Decoding EEG Signals

Ajay Kumar Tanwani<sup>1</sup>, José del. R. Millán<sup>2</sup>, Aude Billard<sup>1</sup>

**Abstract**—Decoding the user intention from non-invasive EEG signals is a challenging problem. In this paper, we study the feasibility of predicting the goal for controlling the robot arm in self-paced reaching movements, i.e., spontaneous movements that do not require an external cue. Our proposed system continuously estimates the goal throughout a trial starting before the movement onset by online classification and generates optimal trajectories for driving the robot arm to the estimated goal. Experiments using EEG signals of one healthy subject (right arm) yield smooth reaching movements of the simulated 7 degrees of freedom KUKA robot arm in planar center-out reaching task with approximately 80% accuracy of reaching the actual goal.

## I. INTRODUCTION

The world around us is going to change markedly with the use of wearable robotic devices assisting humans in everyday tasks. Brain-Machine Interfaces (BMIs) are envisioned to facilitate this integration in the most ‘natural’ way. Decoding brain signals for controlling these devices, however, poses all kind of challenges to existing machine learning and control techniques due to the high-dimensional and non-stationary nature of the data along with the large variability across users. Despite the efforts, there has been a little focus on understanding the high-level *intention* of the user in decoding the brain signals; a fundamental characteristic for practical implementation of such devices.

This paper investigates the use of slow cortical EEG signals in decoding the intention of the user for self-paced reaching movements of a robot arm. *Intention* here refers to an early plan to move that represents a high-level state such as the desired goal to reach as compared to the low-level muscle activations for executing the movement. Contrary to decoding the cue-based movements [1], [2], we consider self-paced reaching movements where the user spontaneously executes the movement without an external cue. Such reaching movements tend to better encapsulate the natural motor behaviour in humans. In this paper, we continuously estimate the current goal/intention throughout the trial starting prior to the movement onset. Previous studies indicate the modulation of slow cortical EEG signals by the intention to move [3]. The decoded goal is used in the reward function to generate optimal trajectories for driving the robot arm to the goal. Our proposed trajectory decoder is easy to learn and generalizes effectively to unseen parts

of the robot workspace. The integrated framework combines the high-level goals encoded in EEG signals with low-level motion plans to control the robot arm in continuous task space. Our target application of this work is to use EEG signals for direct motor control of patients with possibly severe upper-limb disabilities.

## II. MATERIALS & METHODS

### A. Experiments

Experiments were designed to perform center-out planar reaching movements to four goal targets in cardinal directions located 10 cm away from the center, while holding the PHANTOM robotic arm. Four subjects – two healthy and two stroke patients – participated in the experiment carried out at the San Camillo Hospital, Venice, Italy. One patient had left parietic arm with left cerebellar hemorrhagic stroke since 2 months; while other had right parietic arm suffering from left nucleo-capsular stroke since 2 years. After the target was shown to the subject, the subject was asked to wait for at least 2 seconds to perform a self-paced movement (see [3] for details of experimental set-up). For each arm, subjects performed three runs each containing 80 trials each (20 trials per target). Trials were extracted ranging from 2 s before the movement onset until 1 s after the task. For brevity, we only report results of the right arm of the first healthy subject in this work.

The EEG and EOG signals were simultaneously recorded with a portable BioSemi ActiveTwo system using 64 electrodes arranged in an extended 10/20 montage. EOG channels were placed above nasion and below the outer canthi of both eyes in order to capture horizontal and vertical EOG components. The kinematics data of the robotic arm was recorded at 100 Hz, while EEG signals were captured at 2048 Hz and then downsampled to 256 Hz. Preprocessing steps to analyse EEG data required Common Average Referencing (CAR) procedure to remove the global background activity [5]. Moreover, only 34 EEG channels were selected, excluding the peripheral channels and those having high correlation with the EOG activity. EEG signals were then passed through a zero-phase low-pass Butterworth filter with cut-off frequency of 120 Hz, further down-sampled at 128 Hz and finally low-pass filtered at 1 Hz to extract slow cortical potentials. Each EEG channel and kinematic signal was normalized to have zero-mean and unit-standard deviation.

### B. Framework

1) *Intention/Goal Decoder*: To decode the intention/goal in the EEG signals, we perform the online classification in sliding window of 250 ms that shifts by 62.5 ms within the

<sup>1</sup>A. K. Tanwani and A. Billard are with Learning Algorithms and Systems Laboratory (LASA), Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland. <sup>2</sup> J. d. R. Millán is the head of the Defitech Chair on Non-Invasive Brain-Machine Interface (CNBI), Center for Neuroprosthetics in EPFL, Switzerland. {ajay.tanwani, jose.millan, aude.billard} at epfl.ch

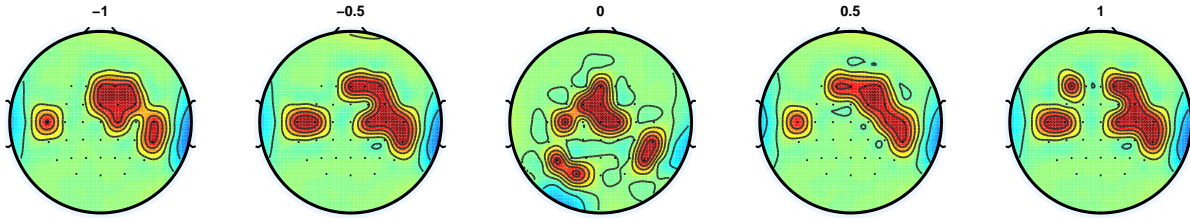


Fig. 1: Evolving EEG channels activity in the time interval  $[-1\ 1]$  seconds

trial period of  $[-2\ 1]$  seconds. Note that we start to decode the goal prior to the movement onset to minimize any delays in controlling the arm (see [3] for details). For each of these windows, the features are selected separately using Canonical Variant Analysis (CVA) with 5 fold cross-validation taking one EEG sample per window at the end. 10 EEG channels with best discriminant power are selected in each window to classify among the 4 target goals. For classification, EEG data is further downsampled to 16 Hz taking into account 4 samples of 10 EEG channels for a total of 40 features. Linear Discriminant Analysis (LDA) [6] is then used for predicting the goal estimate  $x_g$  in every time window from the given EEG feature vector. For the EEG feature vector represented by  $u_t$  at time instant  $t$ , the classification of the goal  $x_{g_t}$  is based on the probability of belonging to each of the goals:

$$x_{g_t} = f(u_t) = \arg \max_{i=1\dots 4} P(C = x_g^{(i)} | u_t) \quad (1)$$

2) *Trajectory Decoder*: The goal of the trajectory decoder is to continuously generate the motion plans to drive the robot arm to the goal. In this paper, we represent this decoder with a dynamical system of the form:

$$\dot{\bar{x}} = f(\bar{x}) + \epsilon \quad (2)$$

where,  $f$  is a continuously differentiable function that maps the 2D-planar Cartesian position of the robot arm  $x$  to its Cartesian velocity  $\dot{x}$ . For simplicity, we transform the coordinates to  $\bar{x} = x - x_g$  to signify the change of all goal positions to the origin of the transformed system. The evolution of robot motion can be computed by integrating Eq. 2. Let  $\alpha \in \mathcal{R}^n$  represent the parameters of the function  $f$ . We are required to learn the parameters  $\alpha$  such that the robot follows the intended movement of the user. To this end, we take a two-step methodology: 1) learn the initial function from demonstrations of the hand kinematics recorded from the subjects using Programming by Demonstration (PbD) [7], and 2) optimize the function parameters for effective generalization using Reinforcement Learning (RL) [8].

In the first stage, we use Support Vector Regression (SVR) to estimate the initial function  $f_i$  given data samples  $\{\bar{x}, \dot{\bar{x}}\}$  from the experiments, represented as:

$$\dot{\bar{x}} = f_i(\bar{x}) = \alpha^T \phi(\bar{x}) + b \quad (3)$$

where,  $\alpha$  represents the weights of the support vectors,  $\phi(\bar{x})$  is the projection of the data  $\bar{x}$  in the  $n$ -dimensional feature

space, and  $b$  is the constant bias. Note that each output dimension is learned separately in this model. To speed up the learning process, we downsample the kinematic data to 5 Hz for a total of 750 samples corresponding to the right arm of the first subject in the training set. Hyper-parameters of the SVR are obtained after grid-search with size of the epsilon-tube,  $\epsilon = 0.5$ , width of the radial basis kernel function  $\gamma = 0.5$ , and complexity parameter  $C = 1$ .

In the second stage, we modify the landscape of the learned function to generate optimal trajectories in the whole state space by maximizing the reward function. The rationale here is to decode the movement effectively far from the training data (see Fig. 4 for clarity). Moreover, optimization in the second stage caters for the imperfection or sub-optimality in the recorded demonstrations (for example, demonstrations of stroke suffering subjects). We express the reward function  $r(\bar{x})$  as:

$$r(\bar{x}) = -w_1 \bar{x}_f^T \bar{x}_f - w_2 \dot{\bar{x}}_f^T \dot{\bar{x}}_f - w_3 \ddot{\bar{x}}_f^T \ddot{\bar{x}}_f \quad (4)$$

where,  $w_1$  weighs the cost for distance from the goal/origin at the end of the trial,  $w_2$  penalizes for any non-zero velocity at the end of the trial, and  $w_3$  is responsible for ensuring smooth movement in reaching the goal by minimizing the norm of the acceleration vector. Weights of the reward function after manual tuning are:  $w_1 = 5$ ,  $w_2 = 0.01$ ,  $w_3 = 0.0001$ . Maximum velocity  $\dot{x}_{max}$  is set to  $30\text{ cm/s}^2$  and the simulations are carried till  $t = 2$  seconds to prolong the penalty by  $w_1$  and  $w_2$  after the end of trial at  $t = 1$  second.

Support vectors of the initial function act as basis functions for the optimized function  $f_o$  in the second stage. Weights of the support vectors  $\alpha$  are optimized by stochastic gradient ascent on the value function,  $J(\bar{x}) = \frac{1}{T} \sum_{t=0}^T r(\bar{x})$ . More precisely, we add noise  $\eta$  sampled from multivariate Gaussian with mean 0 and covariance matrix of  $\sigma^2 I$  with  $\sigma = 0.1$  to the parameters  $\alpha$ , evaluate the value function,  $J(\alpha + \eta)$ , from episodic roll-outs of the current optimized function,  $\dot{\bar{x}} = f_o(\bar{x})$ , and adjust the parameter vector in the direction of increasing value function, i.e.,

$$\Delta \alpha = \beta (J(\alpha + \eta) - J(\alpha)) \quad (5)$$

where,  $\beta$  is a small step-size parameter set to 0.05 in our experiments. The procedure is repeated till the parameter vector stops changing. In our experiments, the parameter vector is improved for 1500 iterations which increases the

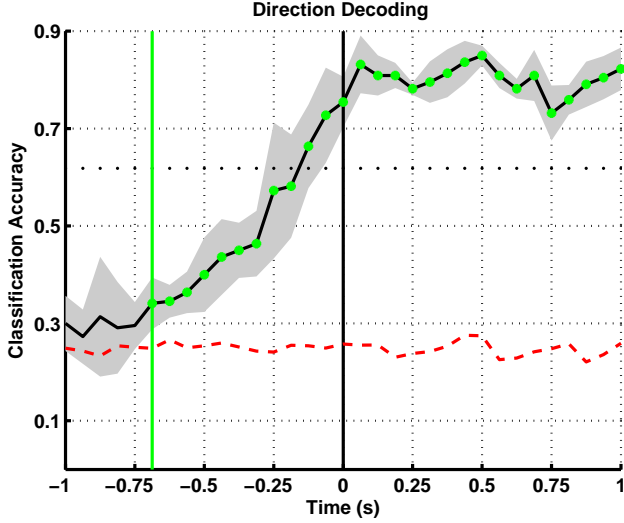


Fig. 2: Decoding goal direction from EEG signals of first healthy subject (right arm). Red line shows the chance level; green line indicates the time instant when the classification accuracy significantly exceeds the chance level; shaded region shows the variation in accuracy over 5-folds.

value of the function parameters  $J(\alpha)$  from  $-118.1$  to  $-4.81$ .

In the proposed framework, the attractor of the optimized dynamical system is shifted from the origin to the estimated goal from Eq. (1) which is updated after every time window of 250 milliseconds. After the end of trial, the optimized dynamical system moves the robot arm to the last estimated goal at  $t = 1$  seconds. Mathematically, the optimized dynamical system takes the form:

$$\dot{x} = f_o(\bar{x} - x_{g_t}) \quad (6)$$

### III. RESULTS

#### A. Decoding Goal

To analyse the performance of the goal decoder from EEG signals, we show the topographic plots of selected channels to depict their discriminatory power at different time instants starting 1 second before the movement onset

TABLE I: Performance comparison of initial and optimized dynamical system using: MSE on the testing set; average correlation in time between simulated and demonstrated position trajectories on the testing set; end-point distance from the goal for different initial conditions

Trajectory Decoder	MSE $cm/s^2$	Correlation [0 1]	End-Point Distance (cm)
Initial SVR	2.49	0.51	5.157
Optimized SVR	-	0.23	0.09

in Fig. 1. As the exact time when movement intent occurs in a self-paced movement is unclear, the plots can provide insights about movement-related modulations in different brain regions during planning and how they evolve over time. It is seen that the activity is dominant in the frontal-parietal regions of brain consistent with earlier reported studies [3].

Fig. 2 reports the classification accuracy of goal decoder in the time window  $[-1, 1]$  seconds. Classification accuracy is computed as the ratio between the sum of correctly classified diagonal entries in the confusion matrix and the total number of instances. The time instant when the classification accuracy significantly exceeds the chance level is used as a

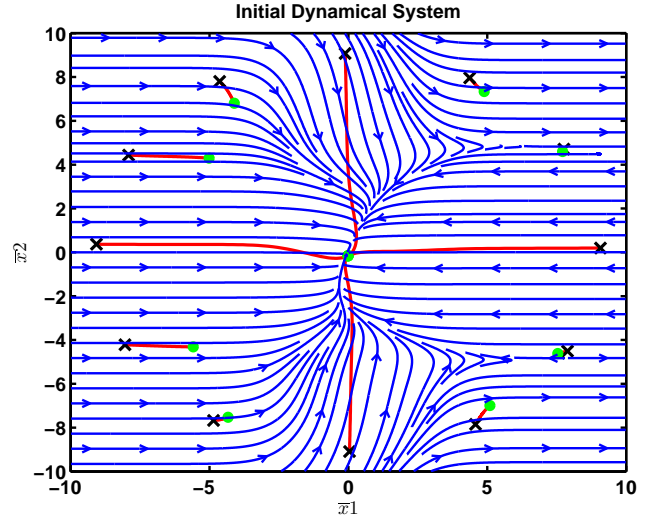


Fig. 3: Performance of initial learned function with SVR. Black crosses indicate the initial positions, while green circles denote the position at the end of the trial

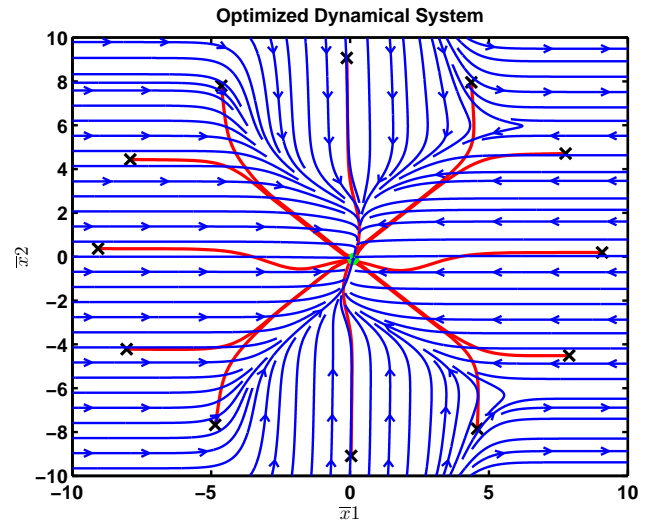


Fig. 4: Performance of optimized function with SVR. Black crosses indicate the initial positions, while green circles denote the position at the end of the trial. Different initial conditions converge to the goal

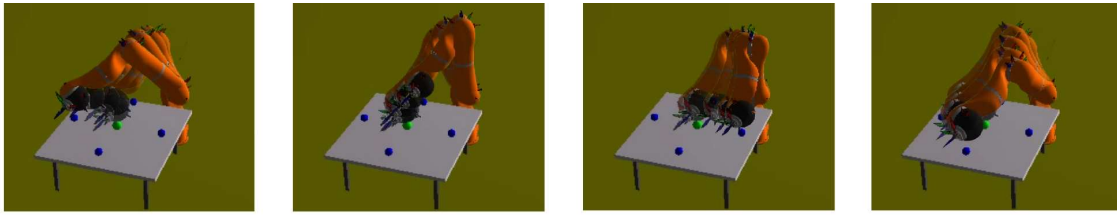


Fig. 5: Simulated trajectories of KUKA robot performing center-out reaching task

metric to initiate the movement with the trajectory decoder. Chance level is calculated by training the classifier on a randomized permutation of the class labels of the training set. Results are then averaged across 10 iterations each with 5 fold cross-validation. Best time for subject 1 is 687.5 ms with classification accuracy of 0.34 before the movement onset (marked with green line in Fig. 2). It is seen that the classification accuracy gradually improves afterwards with a peak accuracy of 0.85 at 0.5 seconds. Further experiments to evaluate the false positive rate, i.e., detecting intention to move when there is actually no movement, is subject to our future work.

### B. Decoding Trajectories

We evaluate the performance of our trajectory decoder using three metrics: 1) Mean-Square Error (MSE) on the training/testing set, 2) Correlation in time of the simulated position trajectories with the demonstrated ones, 3) Distance to the goal at the end of the trial computed by simulating the system from 12 different initial conditions. Table I summarizes the performance of the initial and the optimized SVR. The initial dynamical system learned using SVR performs well in terms of MSE with training and testing error of 2.66 and 2.49  $cm/s^2$  respectively, and a high correlation in position of 0.51 with the demonstrated trajectories. To evaluate the performance of the system far from the training data, we sample 12 different initial points in the plane (shown in Fig. 3 with crosses) and integrate the system forward in time for a period of 2 seconds. As seen in Fig. 3, the initial dynamical system with SVR is not able to generalize away from the training data yielding a high end-point distance error of 5.157 cm. Note that the the initial conditions in the cardinal directions correspond to the training set. On the other hand, optimized SVR is able to drive the robot arm to the goal from all the sampled initial conditions (see Fig. 4). This comes at a cost of relatively low position correlation of 0.23 suggesting the need to further improve the reward function. This generalization is very desirable in our application since the user is expected to control the arm from all parts of the state space.

In Fig. 5, we test the performance of the integrated system on the simulated 7 degrees of freedom KUKA robotic arm. The optimized dynamical system starts to move the robot arm 687.5 milliseconds before the movement onset and finally guides the robot arm to the last estimated goal at the end of the trial. Across all the trials, the robot arm reaches the actual goal with a net accuracy of 79.5% on average. The figure

shows simulated trajectories of the robotic arm reaching different goal positions following the predicted goal from the intention decoder and the optimal motion plans from the trajectory decoder.

## IV. CONCLUSIONS

In this paper, we have presented a system that decodes the intention of the user using non-invasive slow cortical EEG signals and generates optimal motion plans to drive the robot arm to the goal. The most desirable properties of the system include detection of goal direction before the movement onset and generalization of the motion plans away from the training data. In future, we would like to evaluate the performance of the system when the user changes his intention/goal direction during execution of the movement. It will also be interesting to incorporate the EEG signals in our trajectory decoder similar to the works in [9] and [10].

## ACKNOWLEDGEMENT

Authors thank E. Lew, R. Chavarriaga and I. Iturrate for their helpful insights. This work is supported by the National Center of Competence in Research in Robotics (NCCR Robotics).

## REFERENCES

- [1] S. Waldert, H. Preissl, E. Demandt, C. Braun, N. Birbaumer, A. Aertsen et. al. Hand movement direction decoded from MEG and EEG., *Journal of Neuroscience*, 28(4), pp. 1000 - 1008, 2008.
- [2] S. Musallam, B.D. Corneil, B. Greger, H. Scherberger, R. A. Andersen, Cognitive control signals for neural prosthetics., *Science*, 305 (5681), pp. 258 - 262, 2004.
- [3] E. Lew, R. Chavarriaga, S. Silvoni, and J. d. R. Millán, Detection of self-paced reaching movement intention from eeg signals, *Front Neuroeng*, 5 (13), 2012.
- [4] E. Lew, R. Chavarriaga, S. Silvoni, and J. d. R. Millán, Single trial prediction of self-paced reaching directions from EEG signals, *Front Neuroprosthesis*, 2014. *submitted*
- [5] O. Bertrand, F. Perrin, and J. Pernier, A theoretical justification of the average reference in topographic evoked potential studies, *Electroencephalography and Clinical Neurophysiology*, 62, pp. 462 - 464, 1985.
- [6] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, New York: John Wiley, Section, 10 (1), 2001.
- [7] A. Billard, and S. Calinon, and R. Dillmann, and S. Schaal, *Survey: Robot Programming by Demonstration*, *Handbook of Robotics*, chapter 59, 2008.
- [8] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, 1998.
- [9] T. J. Bradberry, R. J. Gentili, and J. L. Contreras-Vidal, Reconstructing three-dimensional hand movements from noninvasive electroencephalographic signals, *J. Neurosci.*, 30(9), pp. 3432 - 3437, 2010.
- [10] N. J. Beuchat, R. Chavarriaga, S. Degallier, and J. d. R. Millán, *Offline Decoding of Upper Limb Muscle Synergies from EEG Slow Cortical Potentials*, 35th Annual Conference on Engineering in Medicine and Biology Society, 2013.