

# A Human-Robot Interface based on Surface Electroencephalographic Sensors

Christos N. Mavridis<sup>†</sup>, John S. Baras<sup>†</sup> and Kostas J. Kyriakopoulos<sup>\*</sup>

**Abstract**—We propose a human-robot interface based on potentials recorded through surface Electroencephalographic sensors, aiming to decode human visual attention into motion in three-dimensional space. Low-frequency components are extracted and processed in real time, and subspace system identification methods are used to derive the optimal, in mean squared sense, linear dynamics generating the position vectors. This results in a human-robot interface that can be used directly in robot teleoperation or as part of a shared-control robotic manipulation scheme, feels natural to the user, and is appropriate for upper extremity amputees, since it requires no limb movement. We validate our methodology by teleoperating a redundant, anthropomorphic robotic arm in real time. The system’s performance outruns similar EMG-based systems, and shows low long-term model drift, indicating no need for frequent model re-training.

## I. INTRODUCTION

Human-Robot Interfaces have been receiving increasing attention in the last decades. In the field of neurorobotics, scientists are focusing on systems that can identify human intention using biosignals in the place of external devices, such as joysticks or game pads. Surface Electromyographic (EMG) signals, generated during muscle contraction [1], have been widely used as control inputs in such systems, as they directly reflect human motion intention [2], [3].

Certain limitations, such as muscle contraction inability and sensor drifting, however, have lead to the use of alternative inputs such as Electroencephalographic (EEG) signals [4]. Muscle activity and motor imagery, have been found to be related with Slow-Cortical Potentials (SCP) and Sensory-Motor Rhythms (SMR) [5], which have been employed in applications such as cursor control [6], virtual helicopter control [7] and hand movement reconstruction [8], [9]. All these applications, however, suffer from low Signal-to-Noise Ratio (SNR), due to the presence of EEG artifacts [10], mainly caused by head and eye movement. Numerous algorithms focus on removing these EEG artifacts, most of them being inappropriate for real-time analysis, and, thus, for robot-control applications. For this reason, experiments have been quite restrictive, often requesting that the subjects keep their eyes closed, avoid blinking, swallowing or any kind of head movement.

<sup>†</sup> Department of Electrical and Computer Engineering and the Institute for Systems Research, University of Maryland, College Park, MD 20742, USA. Emails: {mavridis@umd.edu, baras@umd.edu}. The work of Christos Mavridis and John Baras was partially supported by ONR grant N00014-17-1-2622.

<sup>\*</sup> Control Systems Laboratory, School of Mechanical Engineering, National Technical University of Athens, Greece. Email: kkyria@mail.ntua.gr.

We propose the use of, possibly eye-movement-related, EEG signals in a real-time, continuous human-robot interface that decodes three-dimensional motion. This notion is supported by the use of similar signals in decoding the direction of visual fixation in recent works [11], [12]. The direct use of Electrooculography, however, is limited by large sensor drift and the reduction of the actual field of vision of the subject, and, to date, very few endeavors towards decoding both direction and depth of vision have been made [13].

In this paper, low-frequency components are extracted from potentials recorded over the user’s scalp, processed in real time, and transformed to Cartesian coordinates of the end-effector of a robotic arm in 3D space using a linear time-invariant state space model. This model-based approach is based on numerical methods for subspace-based system identification, and is chosen over recent black-box architectures such as deep neural networks [14] because it allows to be directly used in sensor fusion schemes (e.g. Kalman filtering). This results to a Human-Robot Interface which feels natural to the user, is appropriate for upper extremity amputees –since it requires no limb movement– and can be used in robot teleoperation, or as part of a shared-control robotic manipulation scheme. We validate our methodology by teleoperating a redundant, anthropomorphic robotic arm in real time in three different experiments. In the first experiment, the subjects try natural reaching motions, tracking their hand trajectory with their eyes, in the second experiment, the subjects are asked to follow with their eyes an object which is being moved by a second person, and, in the third experiment, they are seated against the robotic arm, and asked to teleoperate it by looking at the position they want it to go.

The paper is organized as follows: Section II describes the signal processing methods proposed for feature extraction and introduces the theory of numerical subspace-based system identification. Section III explains the hardware and experimental set-up, and Section IV assesses the efficacy of the proposed interface. Finally Section V concludes the paper.

## II. METHODS

In this section, we present our feature extraction methodology and the realization theory for combined deterministic-stochastic, linear time-invariant, state space models. With a slight abuse of definition, we will denote the signals recorded through the surface Encephalographic sensors as EEG signals throughout the paper.

### A. EEG Signal Analysis

We identify the measurements, sampled at a frequency  $f_s = 500\text{Hz}$ , from each of the EEG sensors with a random process  $X_i : \mathbb{N}_0 \times \Omega \rightarrow \mathbb{R}$ ,  $i = 1, \dots, K$ , defined in the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , where  $\mathcal{F} := \{F_n\}_{n \in \mathbb{N}}$  is a filtration with  $F_n = \sigma(X_i(k) | k \leq n)$ , and  $K$  is the number of sensors.

First, an IIR low-pass filter is applied to each  $X_i$ ,  $i = 1, \dots, K$ , in the time domain according to the following difference equation:

$$\hat{X}_i(n) = \frac{1}{a_0} \left( \sum_{j=0}^p b_j X_i(n-j) - \sum_{j=1}^p a_j \hat{X}_i(n-j) \right), \quad n > p, \quad (1)$$

where  $\hat{X}_i$  represents the filtered process for  $i = 1, \dots, K$ , at time instance  $\frac{n}{f_s}$ , and  $a_j$ ,  $b_j$  are the filter coefficients for a Butterworth filter of order  $p = 2$  with cut-off frequency  $f_c = 2\text{Hz}$ . It is important for real-time applications that the filters applied to the signals introduce a small group delay  $\tau_g^* := \max_{\omega} \tau_g(\omega) < 320\text{ms}$ , where  $\tau_g(\omega) = -\frac{d\phi(\omega)}{d\omega}$ ,  $\phi(\omega) = \angle H(j\omega)$ , and  $H(z) = \frac{\sum_{j=0}^p b_j z^{-j}}{\sum_{j=0}^p a_j z^{-j}}$ .

**Remark 1.** A Butterworth filter of order  $p = 2$  with cut-off frequency  $f_c = 2\text{Hz}$ , when applied to a process sampled at  $f_s = 500\text{Hz}$ , introduces a group delay  $\tau_g^* \leq 130\text{ms}$ .

Keeping only the low frequency components of the EEG signals, preserves the ocular information, and absolves them of the fast and non-stationary oscillations. As a result, we make the following assumption

**Assumption 1.**  $\hat{X}_i$ ,  $i = 1, \dots, K$ , are ergodic and stationary random processes.

Next, the mean value of the processes is removed, after being recursively estimated as:

$$E_i(n) = E_i(n-1) + \frac{1}{n} [X_i(n) - E_i(n-1)], \quad n > p \quad (2)$$

such that

$$W_i(n) = \hat{X}_i(n) - E_i(n), \quad n > p, \quad i = 1, \dots, K \quad (3)$$

where  $E_i(p+1) = \hat{X}_i(p+1)$ . Because of the ergodicity and stationarity assumptions, it is easy to see that  $E_i$  converges almost surely to the ensemble average  $\mathbb{E}[\hat{X}_i]$ , as  $n \rightarrow \infty$ , for all  $i = 1, \dots, K$ .

The spatial information and trends carried from the grid of EEG sensors is next exploited by applying a Common-Average Reference (CAR) spatial filter, accentuating components with highly focal distributions [15], as follows:

$$\hat{W}_i(n) = W_i(n) - \frac{1}{K} \sum_{j=1}^K W_j(n), \quad n > p \quad (4)$$

After the spatial filtering, an automatic channel selection algorithm is implemented, based on the Pearson's correlation coefficients

$$\rho_{ij} := \frac{\text{Cov}(\hat{W}_i, Y_j)}{\sigma_{\hat{W}_i} \sigma_{Y_j}}, \quad i \in \{1, \dots, K\}, \quad j \in \{1, \dots, l\} \quad (5)$$

where  $\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$ ,  $\sigma_X = \sqrt{\mathbb{E}[(X - \mathbb{E}[X])^2]}$ , and  $\{Y_i\}_{i=1}^l$  is the set of  $l$  output position features, as defined in Section II-B. The Algorithms reads as follows:

- First two experiments are conducted, and  $\bar{n}$  data samples of  $\hat{W}_i$ ,  $i = 1, \dots, K$ , and  $Y_j$ ,  $j = 1, \dots, l$  are extracted in each experiment;
- $\rho_{ij}$ ,  $i \in \{1, \dots, K\}$ ,  $j \in \{1, \dots, l\}$  are computed for each experiment;
- A channel  $\hat{W}_i$  is selected if  $\rho_{ij} > \bar{\rho}$ , in both experiments, for a threshold value of  $\bar{\rho}$  such that there are at least two channels correlated with each coordinate  $Y_j$ ,  $j = 1, \dots, l$ .

Notice that if  $\mathbb{E}[\hat{W}_i] = \mathbb{E}[Y_j] = 0$ ,  $\forall i \in \{1, \dots, K\}$ ,  $\forall j \in \{1, \dots, l\}$ , then  $\rho_{ij} = \cos \theta_{ij}$ , where  $\theta_{ij}$  is the angle between  $\hat{W}_i$  and  $Y_j$  in the  $\bar{n}$ -dimensional sample space. The locations of the selected channels for Subject B are shown in Fig. 1a.

**Remark 2.** The selected EEG channels are located mainly on the pre-frontal (associated with eye movements) and the occipital (associated with visual processing) lobe areas. In some cases sensors located at the temporal lobe area (associated with head movement and object perception and recognition) were also selected.

At this point, Principal Component Analysis (PCA) [16] is implemented for further dimensional reduction and noise cancellation. The Covariance matrix  $\Sigma$  of  $\hat{W} := [\{\hat{W}_i\}_{i \in I}]$ , where  $I$  is the set of indices selected by the channel selection algorithm and  $[\{A_i\}_{i \in I}]$  represents a matrix whose columns are the vectors  $A_i$ , is pre-computed using samples from the training data and decomposed using Singular Value Decomposition such that:

$$\Sigma = \frac{1}{\bar{n} - 1} \hat{W}^T \hat{W} = V \Lambda V^T \quad (6)$$

where,  $\Lambda$  is a diagonal matrix containing the eigenvalues  $\lambda_i$ ,  $i = 0, \dots, |\mathcal{I}|$  of the symmetric matrix  $\Sigma$ , and  $V$  is an orthogonal matrix ( $V^T = V^{-1}$ ) whose columns are the eigenvectors of  $\Sigma$ . Then the low-dimensional representation  $U := [\{\hat{U}_i\}_{i=0}^m]$  of  $\hat{W}$  is computed in real time by the projection:

$$U = \hat{W}^T V^* \quad (7)$$

where,  $V^*$  is a matrix whose columns are the eigenvectors corresponding to the largest  $m \leq |\mathcal{I}|$  eigenvalues of  $\Sigma$ . The dimension  $m$  is selected such that the first  $m$  principal components describe more than 95% of the total variance.

### B. Kinematic Analysis

The position measurements  $Z_i$ ,  $i = 1, \dots, 3$ , sampled at a frequency  $f_s = 500\text{Hz}$ , correspond to each of the three Cartesian coordinates  $x$ ,  $y$  and  $z$  of the moving target, respectively, measured from the shoulder of the user, as shown in Fig. 1b.

The position signals are filtered with a 3rd-order, low-pass Butterworth filter with cut-off frequency of  $3\text{Hz}$ , following equation (1). Then, the filtered position signals  $\hat{W}_i$ ,  $i = 1, \dots, 3$

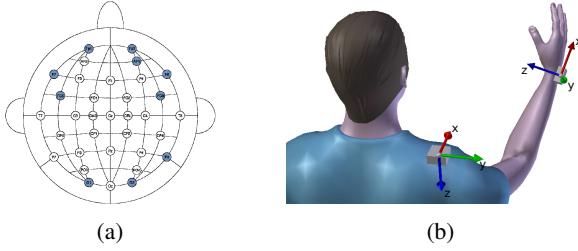


Fig. 1: (a) The locations of the selected EEG channels for Subject *B*. (b) The position tracker sensors.

are mapped within the interval  $[-1, 1]$  using the following transformation  $f : \mathbb{R} \rightarrow [-1, 1]$ :

$$Y_i = f(\hat{Z}_i) := 2 \left( \frac{\hat{Z}_i - Z_i}{\bar{Z}_i - Z_i} \right) - 1, \quad i = 1, \dots, 3, \quad (8)$$

where  $Z_i$ , and  $\bar{Z}_i$  are the minimum and maximum values of  $\hat{Z}_i$ , respectively, pre-computed using the training data. During the testing phase, the reconstructed by the model coordinates  $Y_i$  are projected back to their initial space by solving equation (8) with respect to  $\hat{Z}_i$ ,  $i = 1, \dots, 3$ .

### C. Subspace-based System Identification

We model the decoding from  $u_k = \{\hat{U}_i(k)\}_{i=0}^m$  to  $y_k = \{\hat{Y}_i(k)\}_{i=0}^l$  as a discrete-time, linear, time-invariant, state-space model, defined as:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k \\ y_k &= Cx_k + v_k \end{aligned} \quad (9)$$

where  $x_k \in \mathbb{R}^n$  is a hidden state vector at time instance  $kT$ ,  $n$  is the order of the system,  $T$  is the sampling period,  $u_k \in \mathbb{R}^m$  is the vector of the inputs,  $y_k \in \mathbb{R}^l$  is the output vector, and  $w_k, v_k$  represent zero-mean, Gaussian distributed, white noise, i.e.,  $w_k \sim N(0, Q)$ , and  $v_k \sim N(0, R)$ , with  $\mathbb{E}[wv^T] = S$ .

Given the input and output measurements  $\{u_k\}_{k=1}^N, \{y_k\}_{k=1}^N$ , with  $N \rightarrow \infty$ , and the fact that these two sequences are generated by an unknown model of the form (9), the problem comes down to finding an optimal, in a mean squared error sense, realization  $(A, B, C, D, Q, R, S)$  (up to within a similarity transformation).

This system identification approach is preferred over black-box architectures such as deep neural networks [14] since it can model time dynamics and can be directly used in sensor fusion schemes such as Kalman filtering.

Subspace State-Space Identification (4SID) algorithms have several major advantages in state-space system identification [17]. At first, limited a priori parametrization is needed, i.e. only the order of the system, the state space matrices are not calculated in their canonical forms, which implies that the observability (or controllability) indices do not have to be known in advance, and there is no difference between zero and non-zero initial states. In addition, 4SID algorithms are non-iterative, with no non-linear optimization parts involved, and, as a result, they do not suffer from the

typical disadvantages of iterative algorithms, e.g. not guaranteed convergence, local minima of the objective criterion, and sensitivity to initial estimates.

We make use of the Numerical Algorithm for Subspace State-Space Identification (N4SID) as proposed in [18]. We assume that  $\{A, C\}$  is observable, and  $\{A, BQ^{1/2}\}$  is controllable, and define the extended ( $i > n$ ) observability matrix

$$\Gamma_i = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{i-1} \end{bmatrix} \quad (10)$$

the lower block triangular Toeplitz matrix

$$H_i = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ CA^{i-2}B & CA^{i-3}B & \cdots & D \end{bmatrix} \quad (11)$$

and the ‘past’ and ‘future’ input and output block Hankel matrices  $U_{0|i-1}, U_{i|2i-1}, Y_{0|i-1}$ , and  $Y_{i|2i-1}$ , respectively, with

$$Z_{a|b} = \begin{bmatrix} z(a) & z(a+1) & \cdots & z(a+j-1) \\ z(a+1) & z(a+2) & \cdots & z(a+j) \\ \cdots & \cdots & \cdots & \cdots \\ z(b) & z(b+1) & \cdots & z(b+j-1) \end{bmatrix}, \quad b > a \quad (12)$$

where we presume throughout the paper that  $j \rightarrow \infty$ .

In order to obtain an estimation of  $\Gamma_i$ , we introduce the projection  $Z_i$  of the future outputs onto the past and future inputs and the past outputs, as:

$$Z_i = Y_{i|2i-1} / \begin{bmatrix} U_{0|i-1} \\ U_{i|2i-1} \\ Y_{0|i-1} \end{bmatrix} = [L_i^1 \quad L_i^2 \quad L_i^3] / \begin{bmatrix} U_{0|i-1} \\ U_{i|2i-1} \\ Y_{0|i-1} \end{bmatrix}, \quad (13)$$

where  $A/B = AB^T(BB^T)^{-1}B$ , which corresponds to the optimal prediction of  $Y_{i|2i-1}$  given  $U_{0|i-1}$  and  $Y_{0|i-1}$  [17]. An important observation is that the column space of  $\Gamma_i$ , coincides with the column space of

$$T := [L_i^1 \quad L_i^3] \begin{bmatrix} U_{0|i-1} \\ Y_{0|i-1} \end{bmatrix}, \quad (14)$$

such that, after applying a Singular Value Decomposition

$$T = [U_1 \quad U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} V^T, \quad (15)$$

we can put

$$\Gamma_i = U_1 \Sigma_1^{1/2} \quad (16)$$

and identify  $\Gamma_{i-1}$  as  $\Gamma_i$  with the last  $l$  rows removed.

One can show that

$$Z_i = \Gamma_i \hat{X}_i + H_i U_{i|2i-1} \quad (17)$$

$$Z_{i+1} = \Gamma_{i-1} \hat{X}_{i+1} + H_{i-1} U_{i+1|2i-1} \quad (18)$$

as well as

$$\begin{bmatrix} \hat{X}_{i+1} \\ Y_{i|i} \end{bmatrix} = \begin{bmatrix} A \\ C \end{bmatrix} \hat{X}_i + \begin{bmatrix} B \\ D \end{bmatrix} U_{i|i} + \begin{bmatrix} U_{0|i-1} \\ U_{i|2i-1} \\ Y_{0|i-1} \\ \hat{X}_i \end{bmatrix}^\perp \quad (19)$$

hold, where  $\hat{X}_i$  is the estimation of  $X_i := [x_i, x_{i+1}, \dots, x_{i+j-1}]$ , and  $A^\perp$  indicates a matrix whose row space is perpendicular to the row space of  $A$ .

It follows that

$$\begin{bmatrix} \Gamma_{i-1}^\dagger Z_{i+1} \\ Y_{i|i} \end{bmatrix} = \begin{bmatrix} A & \mathcal{H}_{12} \\ C & \mathcal{H}_{22} \end{bmatrix} \begin{bmatrix} \Gamma_i^\dagger Z_i \\ U_{i|2i-1} \end{bmatrix} + \begin{bmatrix} U_{0|i-1} \\ U_{i|2i-1} \\ Z_i \\ \hat{X}_i \end{bmatrix}^\perp \quad (20)$$

where  $A^\dagger$  denotes the Moore-Penrose pseudo-inverse of  $A$ , and

$$\mathcal{H}_{12} = B - A\Gamma_i^\dagger \begin{bmatrix} D \\ \Gamma_{i-1}B \end{bmatrix} \Gamma_{i-1}^\dagger H_{i-1} - A\Gamma_i^\dagger \begin{bmatrix} 0 \\ H_{i-1} \end{bmatrix} \quad (21)$$

$$\mathcal{H}_{22} = D - C\Gamma_i^\dagger \begin{bmatrix} D \\ \Gamma_{i-1}B \end{bmatrix} - C\Gamma_i^\dagger \begin{bmatrix} 0 \\ H_{i-1} \end{bmatrix} \quad (22)$$

Therefore, by determining the least squares solution:

$$\begin{bmatrix} \Gamma_{i-1}^\dagger Z_{i+1} \\ Y_{i|i} \end{bmatrix} = \begin{bmatrix} \mathcal{H}_{11} & \mathcal{H}_{12} \\ \mathcal{H}_{21} & \mathcal{H}_{22} \end{bmatrix} \begin{bmatrix} \Gamma_i^\dagger Z_i \\ U_{i|2i-1} \end{bmatrix} + \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} \quad (23)$$

where  $\rho_1$  and  $\rho_2$  are residuals, we are in place to

- identify  $A = \mathcal{H}_{11}$ ,  $C = \mathcal{H}_{21}$ ;
- compute  $B$  and  $D$  by solving the linear equations (21) and (22); and
- calculate  $Q$ ,  $R$  and  $S$  by

$$\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} = \frac{1}{j} \begin{bmatrix} \rho_1 \rho_1^T & \rho_1 \rho_2^T \\ \rho_2 \rho_1^T & \rho_2 \rho_2^T \end{bmatrix}. \quad (24)$$

### III. HARDWARE AND EXPERIMENTAL SETUP

Three different experiments were conducted to assess the proposed interface, as shown in Fig. 2. In the first round of experiments, three right-handed subjects (Subject *A*, *B* and *C*), completed a training phase, and Experiments *I* and *II* three times. In the second round, which took place two months after the first, Subject *B* was asked to repeat experiment *I* with no re-training. In addition Subject *B* completed Experiment *I\**, which was a variation of Experiment *I*, as well as Experiment *III*. The duration of the training phases was 60s, while the testing phases lasted for 30s. The best trials of Subject *B* were recorded and are shown in the attached video.

#### A. Experiments *I* and *I\**

The first experiment is designed to test the efficacy of the proposed Human-Robot Interface, compared to existing EMG-based interfaces, and, as a result, involves arm movement. The participants are seated in a chair and instructed to move their right arm freely in 3D space, making natural motions as if trying to reach an object, lift it up, and bring it back towards them (Fig. 2a). At the same time, they are asked to track their moving hand with their eyes in a natural way, allowing smooth motion of their head if needed. No eye-blinking or swallowing restrictions are imposed, such that the interface feels as natural as possible to the user.

EEG signals are recorded over the scalp of the subjects, and their right hand's position is tracked. EMG signals are also recorded from 5 upper-arm muscles and are processed

with a real-time linear envelope method as described in [2] and the references therein. Two different state-space models are trained, using the EEG and the EMG data, respectively. In the testing phase, a robot controller drives the robotic manipulator replicating the motion of the human arm.

In Experiment *I\**, Subject *B* was asked to move their right hand freely in 3D space, while stopping at different times and places, simulating a repeated ‘pick and place’ task. The two different motion profiles are depicted in Fig. 3.

#### B. Experiments *II* and *III*

Experiments *II* and *III* are designed to test the efficacy of the proposed interface when there is no upper limb movement, and, as a result, no EMG data are being used.

In the second experiment (Fig. 2b), the three subjects are asked to track a moving object with their eyes. The object is being moved by a second person, and the robot aims to simulate the movement that the human would do to move the object in the same way.

In the third experiment (Fig. 2c), Subject *B* is placed opposite to the robot, looking directly at the robot’s end-effector, and they are asked to look to the place they want the robot’s end-effector to go.

#### C. Hardware Setup and Robot Control

The EEG signals are recorded from a 32-sensor head cap according to the 10-20 system (Fig. 1a), using a Biosemi ActiveTwo device, and are decimated from  $\hat{f}_s = 16384$  Hz to  $f_s = 500$  Hz using an anti-aliasing, infinite-impulse response (IIR) filter. The EMG signals are recorded with a Delsys Bagnoli device from 5 upper-arm muscles, namely deltoid (anterior), deltoid (posterior), deltoid (middle), pectoralis major (clavicular head) and biceps brachii, which are mainly correlated with the performed motion and less susceptible to noise [1], [2]. The kinematic data are collected using a Polhemus Liberty electromagnetic position tracker at a sample rate of  $\hat{f}_s = 240$ Hz and are interpolated at  $f_s = 500$  Hz. All devices are connected to a personal computer having a TCP communication with the robot’s controller at a frequency of  $f_r = 500$ Hz.

The robot used is a 7-DoF Mitsubishi PA-10 robotic manipulator [19], the first  $d = 4$  DoF of which are controlled such that the robot’s end-effector follows a desired trajectory in 3D space. For the purpose of our experiments, it suffices to use a Jacobian pseudo-inverse controller for redundant manipulators [20]:

$$\dot{q} = K_1 J^+ (Z_d - Z) + K_2 (I_d - J^+ J)(q_r - q) \quad (25)$$

where  $J_{d \times 3}^+ = J^T (JJ^T)^{-1}$  is the Moore-Penrose pseudo-inverse of the corresponding Jacobian matrix  $J$ ,  $K_1$  and  $K_2$  are positive-definite weight matrices,  $Z$  and  $Z_d$  are the actual and desired position of the robot’s end-effector,  $q$  is the vector of the angles of the four robot joints, and  $q_r$  is a reference vector defined in the joint space, such that the robot maintains an elbow-down formation. The desired position vector  $Z = \{[Z_i]_{i=1}^3\}$  is reconstructed by a model of order  $n \in \{2, \dots, 4\}$ . For Experiment *I*, each coordinate  $Z_i$  was



(a) Experiment *I*: The subject moves their hand in 3D space, tracking it with their eyes. The robot replicates the motion.

(b) Experiment *II*: The subject follows a moving object in 3D space. The robotic arm follows the trajectory of the object.

(c) Experiment *III*: The subject is looking at the direction they intend the robot's end-effector to go.

Fig. 2: Experimental Set-Up.

scaled up according to the difference of the length of the human arm and the robotic manipulator.

#### IV. RESULTS

The performance of the learned state-space models is evaluated by two different measures: the root-mean-squared error (RMSE) and Pearson's Correlation Coefficient (CC), defined as:

$$e_{rms,i} = \sqrt{\frac{1}{\bar{n}} \sum_{k=1}^{\bar{n}} (Z_i(k) - \hat{Z}_i(k))^2} \quad (26)$$

$$\rho_i = \frac{\sum_{k=1}^{\bar{n}} (Z_i(k) - \bar{Z}_i)(\hat{Z}_i(k) - \bar{\hat{Z}}_i)}{\sqrt{\sum_{k=1}^{\bar{n}} (Z_i(k) - \bar{Z}_i)^2} \sum_{k=1}^{\bar{n}} (\hat{Z}_i(k) - \bar{\hat{Z}}_i)^2} \quad (27)$$

where  $Z_i$  and  $\hat{Z}_i$ ,  $i = 1, \dots, 3$ , are the actual and reconstructed position trajectories of the robot's end-effector, respectively, and  $\bar{X} = \frac{1}{\bar{n}} \sum_{k=1}^{\bar{n}} X(k)$  denotes the mean value of  $X$  across the  $\bar{n}$  testing samples.

##### A. Comparison with existing EMG-based models

The first two rows of Table I, show the mean and standard deviation of the  $e_{rms}$  and  $\rho$  values of the proposed interface, and the EMG-based interface proposed in [2], across all trials for Experiment *I*. Although tested in limited trials, the proposed methodology suggests greater performance, while requiring no upper limb movement.

##### B. Long-term Model Drift

Model drift refers to a model's predictive performance degrading over time due to a change in the environment that violates the model's assumptions. We test the ability of the proposed interface to maintain its performance for the same user, after a long period of time, with no need for model re-training. During the second round of experiments, which took place two months after the first, Subject *B* was asked to repeat the testing phases of Experiment *I*, for which the state-space models had been trained in the first round. The results are shown in Table II and suggest that the proposed EEG-based Human-Robot Interface, as opposed to the EMG-based, needs no frequent re-training or re-calibration.

##### C. Suitability for Upper-Limb Amputees

The third row of Table I, show the performance of the proposed interface across all trials for Experiment *II*, which shows no statistically significant difference from the performance of the interface in Experiment *I*. This suggests that the actual movement of the arm is not necessary, and, therefore, the proposed interface is suitable for users with upper extremity amputation.

Regarding Experiment *III*, although the system's performance cannot be measured, the subject reported having actual control over the robot arm in terms of it moving in the intended direction.

##### D. Complex Movements and Limitations

The coordinates of the natural human-like movements studied in this work, are, to some extent, correlated, a property that is exploited by the system's architecture. In order to test the accuracy of the learned model during more complex motion profiles, Subject *B* was asked, in Experiment *I\**, to move their right hand freely in 3D space, while stopping at different times and places, simulating a repeated 'pick and place' task. The results are shown in the last row of Table I, and the two different motion profiles, along with the reconstructed by the model trajectories for a random trial of Subject *B*, are depicted in Fig. 3.

Although tested in few testing trials, the system's performance metrics show a small decrease in the accuracy of the interface, which still remains, however, at state-of-the-art levels. More complex motion profiles may negatively affect the overall performance of the proposed interface and especially the reconstruction of the depth ( $x$ -coordinate). Further experimentation is needed to quantify this effect.

#### V. CONCLUSION

A human-robot interface was proposed, based on potentials recorded through surface Electroencephalographic sensors, in order to decode human visual attention into position in space. Low-frequency potentials were used in a subspace system identification algorithm to identify the optimal, in mean squared sense, linear time-invariant system generating a three-dimensional position vector in real-time. The resulting human-robot interface feels natural to the user, requires no limb movement, and can be used in robot teleoperation, or as part of a shared-control robotic manipulation

Input Signals	Data	$CC_x$	$CC_y$	$CC_z$	$RMSE_x(cm)$	$RMSE_y(cm)$	$RMSE_z(cm)$
EMG	Experiment <i>I</i>	$0.79 \pm 0.05$	$0.84 \pm 0.04$	$0.87 \pm 0.05$	$6.0 \pm 1.09$	$7.1 \pm 0.53$	$7.9 \pm 0.94$
EEG	Experiment <i>I</i>	$0.85 \pm 0.06$	$0.93 \pm 0.03$	$0.89 \pm 0.05$	$3.34 \pm 0.39$	$5.49 \pm 1.81$	$6.56 \pm 2.2$
EEG	Experiment <i>II</i>	$0.83 \pm 0.07$	$0.92 \pm 0.05$	$0.90 \pm 0.04$	$3.64 \pm 0.45$	$6.38 \pm 2.01$	$6.32 \pm 2.16$
EEG	Experiment <i>I*</i>	$0.78 \pm 0.04$	$0.95 \pm 0.01$	$0.79 \pm 0.05$	$2.97 \pm 0.31$	$5.3 \pm 1.26$	$8.02 \pm 1.17$

TABLE I: The decoding accuracy of the model in terms of the mean and standard deviation values of  $e_{rms}$  and  $\rho$  across all trials and subjects for every experiment.

Input Signals	$CC_x$	$CC_y$	$CC_z$	$RMSE_x(cm)$	$RMSE_y(cm)$	$RMSE_z(cm)$
EMG	$0.24 \pm 0.22$	$0.64 \pm 0.16$	$0.77 \pm 0.11$	$10.9 \pm 1.33$	$11.7 \pm 1.08$	$9.3 \pm 1.12$
EEG	$0.81 \pm 0.04$	$0.93 \pm 0.08$	$0.88 \pm 0.07$	$4.9 \pm 0.98$	$6.6 \pm 1.24$	$5.5 \pm 1.55$

TABLE II: The decoding accuracy for Subject *B* in Experiment *I*, two months after the model training.

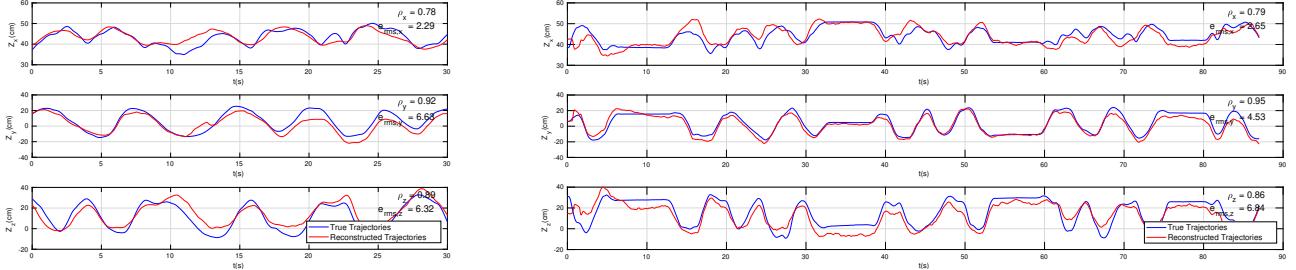


Fig. 3: The actual and decoded trajectories for a random trial of Subject *B* in Experiment *I* (left) and Experiment *I\** (right).

scheme, even by upper extremity amputees. We validated our methodology by teleoperating a redundant, anthropomorphic robotic arm in three different real-time experiments. The system's performance outruns similar EMG-based systems, and shows no need for frequent re-training for each user.

The proposed interface can potentially be combined with a virtual reality headset in order to teleoperate a robot in inaccessible for humans environments, or enhance the experience of virtual computer games.

## REFERENCES

- [1] J. Cram and G. Kasman, *Introduction to Surface Electromyography*. MD: Aspen, Gaithersburg 1998.
- [2] P. Artemiadis and K. Kyriakopoulos, “Emg-based control of a robot arm using low-dimensional embeddings,” *IEEE Transactions on Robotics*, vol. 26, no. 2, pp. 393–398, 2010.
- [3] M. Ison and P. Artemiadis, “The role of muscle synergies in myoelectric control: Trends and challenges for simultaneous multifunction control,” *Journal of Neural Engineering*, vol. 11, no. 5, 2014.
- [4] J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, and T. Vaughan, “Brain-computer interfaces for communication and control,” *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002.
- [5] H. Yuan and B. He, “Brain-computer interfaces using sensorimotor rhythms: Current state and future perspectives,” *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 5, pp. 1425–1435, 2014.
- [6] J. Wolpaw and D. McFarland, “Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 51, pp. 17849–17854, 2004.
- [7] A. Doud, J. Lucas, M. Pisansky, and B. He, “Continuous three-dimensional control of a virtual helicopter using a motor imagery based brain-computer interface,” *PLoS ONE*, vol. 6, no. 10, 2011.
- [8] T. Bradberry, R. Gentili, and J. Contreras-Vidal, “Reconstructing three-dimensional hand movements from noninvasive electroencephalographic signals,” *Journal of Neuroscience*, vol. 30, no. 9, pp. 3432–3437, 2010.
- [9] N. Robinson, A. Vinod, and C. Guan, “Hand movement trajectory reconstruction from eeg for brain-computer interface systems,” 2013, pp. 3127–3132.
- [10] G. Philips, M. Hazrati, J. Daly, and J. Principe, “Addressing low frequency movement artifacts in eeg signal recorded during center-out reaching tasks,” *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, vol. 2014, pp. 6497–6500, 2014.
- [11] A. Belkacem, D. Shin, H. Kambara, N. Yoshimura, and Y. Koike, “Online classification algorithm for eye-movement-based communication systems using two temporal eeg sensors,” *Biomedical Signal Processing and Control*, vol. 16, pp. 40–47, 2015.
- [12] R. Ramli, H. Arof, F. Ibrahim, M. Idris, and A. Khairuddin, “Classification of eyelid position and eyeball movement using eeg signals,” *Malaysian Journal of Computer Science*, vol. 28, no. 1, pp. 28–45, 2015.
- [13] C. Stevenson, T.-P. Jung, and G. Cauwenberghs, “Estimating direction and depth of visual fixation using electrooculography,” vol. 2015–November, 2015, pp. 841–844.
- [14] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, “Deep learning with convolutional neural networks for eeg decoding and visualization,” *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.
- [15] D. McFarland, L. McCane, S. David, and J. Wolpaw, “Spatial filter selection for eeg-based communication,” *Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 3, pp. 386–394, 1997.
- [16] K. Pearson, “On lines and planes of closest fit to systems of points in space,” *Philosophical Magazine*, vol. 2, pp. 559–572, 1901.
- [17] P. Van Overschee and B. De Moor, *Subspace identification for linear systems : theory, implementation, applications*. Boston: Kluwer Academic publ, 1996. [Online]. Available: <http://opac.inria.fr/record=b1077808>
- [18] P. V. Overschee and B. D. Moor, “N4sid: Subspace algorithms for the identification of combined deterministic-stochastic systems,” *Automatica*, vol. 30, no. 1, pp. 75 – 93, 1994, special issue on statistical signal processing and control. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0005109894902305>
- [19] N. A. Bompos, P. K. Artemiadis, A. S. Oikonomopoulos, and K. J. Kyriakopoulos, “Modeling, full identification and control of the mitsubishi pa-10 robot arm,” in *Advanced intelligent mechatronics, 2007 IEEE/ASME international conference on*. IEEE, 2007, pp. 1–6.
- [20] L. Sciavicco and B. Siciliano, *Modeling and Control of Robot Manipulators*. McGraw-Hill, New York, 1996.