Guess What I Attend: Interface-free Object Selection Using Brain Signals

Henrich Kolkhorst[1], Michael Tangermann[2] and Wolfram Burgard

***Abstract*— Interpreting the brain activity to identify user goals or to ground a robot's hypotheses about them is a promising direction for non-intrusive and intuitive communication. Such a capability can be of particular relevance in the context of human-robot cooperation scenarios. This paper proposes a novel approach to utilize the natural brain responses to highlighted objects in the scene for object selection. By this, it circumvents the need for additional interfaces or user training. Our approach uses methods from information geometry to classify the target/non-target response of these event-related potentials. Online experiments carried out with a real robot demonstrate an accurate detection of target objects solely based on the user's attention.**

I. Introduction

When assisting and interacting with humans, it is beneficial if robots can infer the user's intent and comply with it. While in certain situations user goals can be inferred well from context or preferences, users sometimes need the possibility to give explicit commands or to choose from a set of potential goals. One example is the selection of an object from a scene that a robotic assistant should fetch and deliver. Here, the human could use different modalities to select objects, ranging from screen-based interfaces over speech and gaze control to brain-computer interfaces (BCIs). Specifically for impaired users that have problems to reliably exert control over muscular pathways, BCI systems offer an additional control channel to interact with the environment.

In this paper, we address the problem of selecting objects using a non-invasive BCI. In our setting, a robot triggers the (passive) human brain response upon highlighting different objects *in the scene* by decoding the corresponding brain responses from the electroencephalography (EEG).

We leverage the ability of robots to interact with the environment for direct object selection, avoiding the use of an additional interface. This promises to reduce the user's workload. Additionally, using the spatial position of objects avoids the need to disambiguate between visually similar objects and the reference resolution needed in speech-based approaches. While gaze-based interfaces follow a similar approach, they suffer from the problem of having to distinguish "aimless" gaze from gazing at a target object purposefully in order to select it.

To avoid these problems, we propose to select target objects by having the robot gesture towards objects. We implement this by highlighting an object with a laser pointer. This only occupies a small portion of the subject's field-of-view while still leveraging the user's *attention*.

All authors are with the Department of Computer Science, University of Freiburg, Germany. This research was supported by the German Research Foundation (DFG, grant number EXC 1086) and by the state of Baden-Württemberg, Germany, through bwHPC (grant INST 39/963-1 FUGG).

[1] kolkhorst@informatik.uni-freiburg.de
[2] michael.tangermann@blbt.uni-freiburg.de

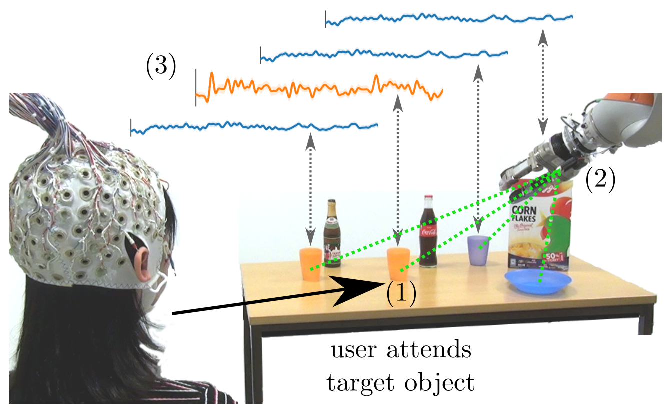

Fig. 1. Selection of objects by attending them: The user attends to a target object (1). The robot probes for the target object by highlighting candidate objects in the scene (2). Based on the difference of the evoked response for target and non-target objects in the brain signals recorded using EEG, the robot infers the target object (3).

In the following, we present the details of an interface-free BCI-based approach to object selection. In summary, we make the following contributions in this work: (1) an EEG-based framework that lets the user select scene objects by attending them, including a non-intrusive highlighting strategy to evoke a brain response in the user without requiring an additional interface; (2) an online processing and classification pipeline based on information geometry that effectively discriminates between target and non-target responses; and (3) experimental results with six untrained users and a real robot in which target objects are successfully inferred in 93 % of online trials.

II. Related Work

As an elementary part of human-robot interaction, a wide range of approaches to target object selection have been proposed. [1]. In this section, we focus on selection mechanism using BCIs and their applications in robotics.

In order to extract information from a user's brain signal, literature on BCI suggests different well-suited experimental paradigms. Besides utilizing self-driven imagery tasks (e.g., the imagination or attempted execution of different movement classes), many BCI setups exploit external stimuli. Visual, auditory or somatosensory stimuli presented to a human subject elicit a transient potential in the brain signals of the observer and can be observed in the EEG. Interestingly, attention modulates the average timing characteristics, amplitudes and the spatial location characteristics of these so-called event-related potentials (ERP).

Despite the high noise level in typical brain signal recordings, BCI systems are capable of analyzing the ongoing brain signals by classifying ERP responses evoked by attended *target* stimuli from those of ignored *non-target* stimuli. State-of-the-art approaches make use of shrinkage-regularized Linear Discriminant Analysis (LDA) classification models which are trained to distinguish target- from non-target ERP responses [2].

Recently, alternative approaches for ERP classification in BCI have been proposed by Barachant et al. Instead of using EEG potentials for classification, they propose to represent each EEG epoch as a covariance matrix and realize the classification in the space of covariance matrices using information geometry to measure distances between these matrices [3], [4].

BCI applications based on ERP paradigms can allow motor-impaired users to choose between multiple options, each represented by a specific stimulus. To make a choice, users focus their attention for example on one out of several symbols on a computer screen while all symbols are visually highlighted multiple times in a rapid stimulation sequence. A single selection process, which we call a *trial*, typically involves multiple repetitions of all symbol highlightings. By analyzing an increasing number of segments of EEG activity (*epochs*), the BCI collects evidence for the target symbol over the course of the trial. ERP-based BCI systems have been successfully used for, e.g., text spelling [5], media applications [6] and gaming [7]. Using such systems is straightforward for software applications that can be realized by symbols on a computer screen. If the device to be controlled, however, is physical (e.g., a wheelchair), the cognitive workload for the user induced by switching between the device and its control via a screen can be substantial.

Several publications in recent years have presented applications of BCIs in robotics. While invasive measurements using intracortical arrays enable direct control of robots [8], the challenging signal-to-noise ratio in non-invasive BCIs suggest the use of a small number of high-level commands or nested decision interfaces. Using a traditional screen-based control interface, BCI spellers have been adapted for action selection in a wheelchair setting [9]. Other applications such as grasp selection involve the use of a screen interface showing potential commands in rapid succession [10]. Movement intention has also been used to decode the user target for a robotic arm in a simulated environment [11].

Mental tasks (such as hand movement, foot movement or mental rotation) are an alternative command strategy. In an integrated setup with a mobile manipulation robot, Burget and colleagues decoded user commands in response to a screen-based interface [12], for which mental imagery is used. The number of decodable mental tasks is limited and strongly varies between users, which is important in practical application. Additionally, the mental task is typically

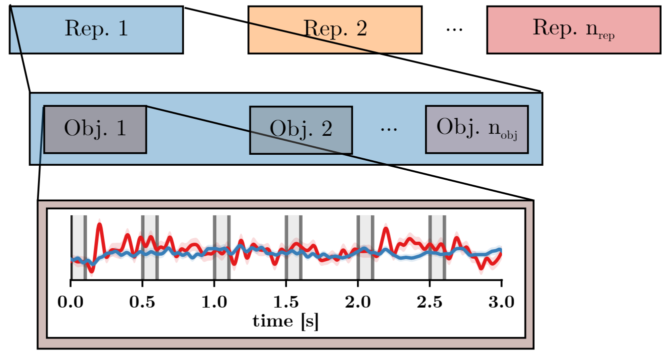

Fig. 2. Course of one experiment trial (in which the target object remains constant): Objects are highlighted every $0.50\,\mathrm{s}$ over a period of $3\,\mathrm{s}$ (bottom row). This is performed for every candidate object (center row). In order to improve object identification and reduce the influence of artifactual epochs, these so-called repetitions are performed multiple times (top row).

unrelated to the actual user goal (e.g., selecting a cup or a plate), and hence can be less intuitive and result in a higher user workload than the proposed attention-based interface.

III. BCI-based Approach to Target Selection

Identifying the user's target object in a scene using an ERP-based BCI requires three main components: First, a decodable brain response must be elicited by appropriate stimuli in an experimental paradigm. The corresponding responses must be classified on an individual level, i.e., into target and non-target classes, and subsequently these individual noisy estimates have to be aggregated to be useful in the actual task, in our case object selection.

A. Object Highlighting to Elicit Visual Evoked Potentials

In pursuit of identifying the user's target object, we highlight relevant candidate objects in order to evoke an event-related response. Highlighting is performed using a laser pointer, which allows for both an accurate marking of the object (also outside of the robotic arm's workspace) and for a reduced ambiguity compared to, e.g., pointing with the robotic hand's finger. In addition, this way of proposing objects allows to deliver the stimuli with a precise timing which eases the analysis of the user's brain signals.

As depicted in Fig. 2, we illuminate each object multiple times per trial. While most experimental paradigms for ERPs realize a pseudo-random sequence of target-/non-target stimuli, this would be impractical in our scenario, as a switch between objects after each single highlighting event incurs the cost of having to move the robotic arm. Delivering all highlighting stimuli for an object consecutively would minimize the movements of the arm at the cost of high interdependence between observed EEG epochs. As a compromise, we chose to highlight each object consecutively for six times over a period of $3\,\mathrm{s}$ before switching to the next object, and repeat this for every object n_{rep} times.

B. Decoding Target Responses from EEG recordings

Due to the large inter-subject variability in brain signals recorded with EEG, we perform a subject-specific calibration phase, in which labeled data is gathered for classifier training. Following novel approaches by Barachant and colleagues [13], [4], we utilize the covariance structure provided by each single EEG epoch for classification. Note that epoch-wise covariance matrices are more noisy than class-wise covariance matrices that are the basis for state-of-the-art ERP classification with LDA [2].

1) Data Representation: We segment the EEG recording into *epochs* $X_i \in \mathbb{R}^{n_c \times n_s}$ of fixed lengths that are time-aligned to each object illumination, where n_c is the number of channels and n_s the number of time samples in each epoch. Since the temporal structure of each time course is critical for the classification of event-related potentials, we augment each epoch with *template prototype signals* P^+, P^- of the class-specific responses (target and non-target responses in our case).

$$\tilde{X}_i = \begin{pmatrix} P^+ \\ P^- \\ X_i \end{pmatrix} \tag{1}$$

Class-wise means of all recorded channels on the training data would be an obvious choice for the template. However, this would triple the size of each epoch and would therefore result in a nine-fold increase in the entries of the covariance matrix. Hence, we resort to dimensionality reduction techniques for projecting the class-wise response to lower-dimensional time series. In this work, we use the xDAWN algorithm [14] to select three surrogate channels per class (i.e., $P^+, P^- \in \mathbb{R}^{3 \times n_s}$). We denote the number of rows in $\tilde{X}_i$ with n'_c. Since the estimation of such projections in xDAWN requires access to class labels, we estimate them on the training data. Note that unsupervised dimensionality reduction methods are also possible. For each epoch i, we calculate the regularized covariance C_i of $\tilde{X}_i$ with an analytically determined shrinkage (Ledoit-Wolf).

2) Riemannian tangent space transformation: Since the Euclidean distance in the space of covariances has undesirable properties [15], yet common classifiers expect an Euclidean feature space, we utilize that covariance matrices are members of the space of symmetric positive-definite (SPD) matrices $\mathcal{P}_n = \{X \in \mathbb{R}^{n \times n} \mid X = X^T, X \succ 0\}$. We project each epoch covariance into the tangent space of the manifold of SPD matrices at a reference point C_{ref}. Following [3], we do this using a logarithmic mapping:

$$\tilde{S}_i = \operatorname{logm}\left(C_{\text{ref}}^{-1/2} C_i C_{\text{ref}}^{-1/2}\right) \tag{2}$$

Here, logm corresponds to the logarithm of a diagonalizable matrix (i.e., the logarithm of each element of the diagonal after the corresponding decomposition is taken).

Let $\|\cdot\|_F$ denote the Frobenius norm. In order for the Euclidean distance in the tangent space to approximate the geodesic distance between two SPD matrices $\delta_R(C_1, C_2) = \|\log\left(C_1^{-1} C_2\right)\|_F$, the covariances C_i to be projected should be distributed closely around the reference point C_{ref} in Equation 2. Following [13], we choose the Frechét mean (also known as geometric mean) of the covariance matrices in the training data for this purpose:

$$C_{\text{ref}} \in \underset{C \in \mathcal{P}_{n'_c}}{\arg\min} \sum_{1 \le i \le n_{\text{train}}} \delta_R^2\left(C, C_i\right) \tag{3}$$

Having transformed each augmented epoch $\tilde{X}_i$ into the tangent space representation $\tilde{S}_i \in \mathrm{R}^{n'_c \times n'_c}$, we keep only the upper triangular part (multiplied by $\sqrt{2}$ for all off-diagonal elements) and vectorize it, yielding a $(n'_c)(n'_c + 1)/2$-dimensional feature vector $\tilde{\mathbf{s}}_i$.

3) Discriminating between Target and Non-Target Responses: In order to distinguish for an epoch i whether it is a response to a target or non-target stimulus, we train a Logistic Regression classifier with $\mathcal{L}_2$ penalty based on the vectorized versions of $\tilde{S}_i$. Thus we get an approximation $P(\text{is\_target} \mid \tilde{\mathbf{s}}_i) = \sigma(\mathbf{w}^T \tilde{\mathbf{s}}_i)$ of the likelihood of epoch i being a response to a target object. We denote the results of this classifier with TCOV in the following section.

We compare our results against a best practice approach from [2] that uses mean voltages in n_{int} intervals of $50\,\text{ms}$ lengths at all channels as features, thus leading to $n_c \cdot n_{\text{int}}$-dimensional feature vectors. Discrimination between target and non-target epochs is performed using LDA for which the covariance matrices are regularized with an analytically determined shrinkage. We denote this approach with IVAL in the results.

C. Aggregating Stimulus Responses to Select Target Objects

For each trial k, we obtain $6 \cdot n_{\text{obj}} \cdot n_{\text{rep}}$ epochs that are aligned to the highlighting stimuli (c.f., Fig. 2). We denote the indices of all epochs of trial k in which object o_j was highlighted with $\mathcal{I}_{k,j}$.

Under the simplifying assumption that epochs, i.e. highlighting responses, are independent of each other, we predict j^* as the user target such that

$$j^* \in \arg\max_j \prod_{i \in \mathcal{I}_{k,j}} P(\text{is\_target} \mid \tilde{\mathbf{s}}_i) \tag{4}$$

where the latter probability that an epoch corresponds to a target stimulus is estimated by the binary classifier.

IV. EXPERIMENTS

A. Setup and Data Collection

We report results from six healthy subjects, who each participated in a single session of an online experiment. Following the declaration of Helsinki, we received approval by the local ethics committee for this study and obtained written informed consent from participants prior to the session.

We seated the subjects approximately $1.20\,\text{m}$ in front of a table with objects. Using a focused laser beam, the visually highlighted sections on the objects covered only small fractions of the subject's field of view. We instructed the subjects to put themselves in the condition of a user of an assistive robotic arm and that they could decide which object the robot should fetch by attending the object. In order to support the performance evaluation, we determined the target object according to the experimental protocol and gave it as a cue to the subject prior to the start of every trial.

For the experiments, we assumed identical prior probabilities for all objects. With each subject, we performed 48 trials with four objects on the table. Note that objects do not have to be modified and that we include visually identical objects (c.f., Fig. 1). For subjects S5 and S6, twelve additional trials were performed with seven objects to check how well the approach can generalize to a different object count.

Experiments were performed with a Kuka iiwa robotic arm with seven degrees of freedom. It is however applicable to any robotic arm that can accurately reach a position to highlight a desired object. The laser pointer was mounted next to the end effector. It was activated by an Arduino microcontroller which also sent time markers to the EEG recording hardware to mark the start of each laser illumination. These markers allowed to segment the EEG data into epochs during online and offline analyses.

Every object was highlighted from an object-specific end effector pose which was determined as a trade-off between maximizing the visibility by the participant and minimizing the robot's trajectories lengths for switching between poses. An example of an experimental trial can be found in the accompanying video.

The course of an experimental session is depicted in Fig. 2. For the online experiment, we chose $n_{\text{rep}} = 3$ and $n_{\text{train}} = 9$ training trials. We also report results for varying theses parameters using "simulated online" processing, for which the only difference is that pre-recorded EEG signals are used. For a whole session consisting of 48 trials, we recorded 2,592 non-target and 864 target responses. The robotic arm took on average $0.71 \pm 0.15\,\text{s}$ to switch between the highlighting poses. In addition, we added $0.10\,\text{s}$ of waiting time before and after each robot move to reduce the influence of the sound of the robot's servo motors upon the adjacent epochs.

We acquired the brain signals using a cap holding 32 Ag/AgCl gel-based passive EEG electrodes positioned according to the extended 10-20 system with a nose reference. Channel impedances were kept below $20\,\text{k}\Omega$. The amplifier sampled the EEG signals at $1\,\text{kHz}$. For online analysis, we used only a subset of channels known to be discriminative for visual ERP responses (channels T7, Cz, T8, P7, P3, Pz, P4, P8, O1, O2).

We streamed the raw EEG signals via a local network to the control computer running ROS. For parts of the EEG analysis, we utilized the MNE-Python [16] and pyRiemann [4] software packages. We filtered data to a band of $0.50\,\text{Hz}$ to $16\,\text{Hz}$ using a FIR filter before downsampling to $100\,\text{Hz}$. We extracted epochs from $-0.20\,\text{s}$ to $1.00\,\text{s}$ relative to the stimulation onset leading to $n_s = 121$ samples per epoch and subsequently corrected them separately for signal drifts using the first $0.20\,\text{s}$ as a baseline. To ensure observations for each object, no rejection was performed in the online setup. For the visualizations in Fig 3, we rejected epochs in which the peak-to-peak amplitude exceeded $100\,\mu\text{V}$. We performed the "simulated online" experiments for the baseline condition (IVAL) with identical processing (e.g., the same causal filtering).

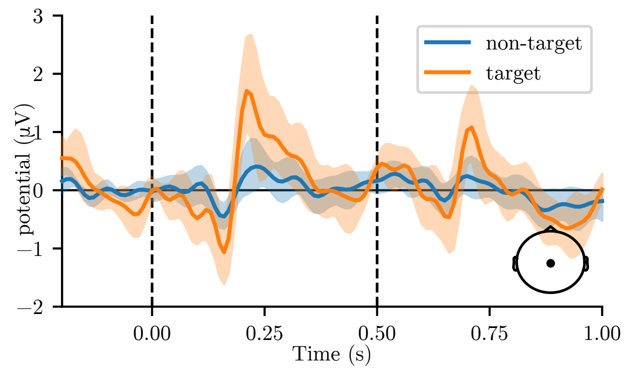

Fig. 3. Grand average mean responses to target (red) and non-target (blue) stimuli for all six subjects recorded at electrode Cz (location depicted in the bottom right of the plot). Dashed lines mark the onsets of two object highlighting events. Shaded blue and red areas mark the corresponding 95th percentile range determined by a bootstrap test.

B. Results

After validating the brain response elicited by the object highlighting, we evaluate our proposed attention-based object selection both regarding the classification performance on individual epochs as well as the correct decision after a trial.

1) Electrophysiology of Response: While online data analysis was performed on single epochs, we conducted an offline analysis of the collected data to visualize the event-related responses. The grand averages of target and non-target responses over all subjects are depicted in Fig. 3. Laser illumination of an attended target object resulted in a more negative deflection at the channel Cz approximately $180\,\text{ms}$ after the stimulus onset compared to responses evoked by non-target stimuli. In addition, attended target stimuli evoke a stronger positive deflection in the period of $200\,\text{ms}$ to $350\,\text{ms}$ after stimulus onset, commonly referred to as the P300 component.

2) Target Decoding Performance: We evaluate the classification performance for single epochs into targets and non-targets using the area under the receiver-operating characteristic (AUC). Training and test data is split chronologically to avoid training on closely correlated training samples. After training on data collected during a calibration period of 9 trials (see training errors listed in Table I), the proposed classification using Riemannian geometry (TCOV) achieves a mean AUC value of 0.80 during the online session. This outperformed the classification based on interval means which achieves a mean AUC of 0.72.

Overall it would be desirable, if the system can be set up using a minimum amount of data for training and online decisions. By discarding complete stimulus repetitions from the data and by changing the number of trials used for classifier training, we can investigate the influence of these parameters on system performance. Since epochs from each repetition are separated in time, we assume that the brain response is not altered due to changing n_{rep}.

The top row of Fig. 4 depicts the improvements in

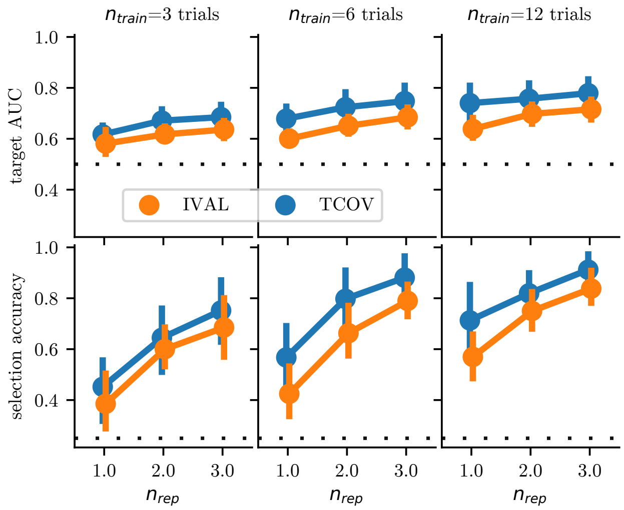

Fig. 4. Decoding quality of single epochs (target/non-target) (top row) and object selection accuracy (bottom row) for varying training sizes (columns) and repetitions (x-axis values). Results for $n_{\text{train}} = 9$ can be found in Table I. Trial selection accuracy is reported for $n_o = 4$ objects. The point estimates correspond to the mean over 6 subjects, with the error bars corresponding to a bootstrapped $95\,\%$ confidence interval. The dotted horizontal lines depict the theoretical chance level.

target/non-target classification performance which can be achieved with varying amounts of training data. While the difference is smaller in the regime with few training data, the classification of augmented covariance matrices (TCOV) outperforms the IVAL approach consistently. Note that the change in AUC value with varying repetitions is due to differing amounts of training data, since we fixed the number of training trials, but not of repetitions.

In order to get an estimate of the performance in a regime with larger amounts of calibration data, as would be expected for a real user, we list the performance on a 5-fold chronological cross-validation (CV) in Table I. One can note that the online performance that we report is still limited by the amount of calibration data and not the inherent noise in the EEG signal. With increasing amounts of training data, AUC values improve by approximately $10\,\%$, to a mean value of 0.84 for the cross-validation in the TCOV case.

3) Object Selection Performance: Moving from the stimulus-specific binary classification to the object-selection task, we evaluate the selection accuracy, i.e., the fraction of trials in which the target object was predicted correctly. In the online experiments, where subjects performed 39 test trials with four objects each, the target object was correctly identified in $93\,\%$ of trials, with two out of six subjects having each object selection correct (c.f., Table I). The prediction using TCOV performed significantly better than IVAL for each subject (Wilcoxon signed-rank test, $p < 0.05$).

Identification of target objects obviously depends on the quality of the underlying target classification and therefore varies with the amount of training data (c.f., Fig. 4). However, the number of repetitions n_{rep}, i.e., the amount of evidence collected for each object, is critical for performance as well. For all three calibration phase lengths, the selection performance improves substantially with each additional repetition. While only three repetitions were performed in the reported experiments, pilot measurements with four repetitions showed only minor additional improvements in selection accuracy for the fourth repetition. In the additional trials with seven objects, which we performed with subject S5 and S6, the robot selected the correct object in $70\,\%$ of trials, with an expected drop in performance due to the larger amount of candidate objects without an increase of evidence for the target.

V. Discussion

The online experiments show that target objects can reliably be identified based on the brain response to the object highlighting in the scene. The proposed new BCI interaction scenario has a number of advantages compared to object selection using other modalities: First, the control mode is intuitive and does not require any user training.

Second, it does not create additional workload as the objects are probed directly. With the presented approach, selecting an object neither requires the user to map a command to a differing target action nor unnecessary cognitive effort as the user can continuously keep his attention on the scene. In questionnaires answered after the experiment session, users rated the task as being easy and stated that they felt relaxed. Screen-based interfaces on the other hand can create a cognitive mismatch between the physical object's location and the spatially separated screen, and visually similar objects could be mixed up. Selecting objects by speech commands would require reference resolution and natural language understanding.

Third, compared to object selection via an additional screen or compared to speech commands, the selection of an object with this BCI is non-intrusive with respect to other tasks. Considering users that have remaining physical capabilities, instructing the robot will rarely be their main task but rather a secondary aspect of some primary task. Imagining a household robot serving food or a robot assisting a surgeon, an important aspect of interaction is how intuitive and non-intrusive the instruction is. Language commands can be intrusive in social situations, while using a laser-based selection the user could even be engaged in another task.

Using a BCI, however, requires the user to wear an EEG cap. Its setup—at least for current gel-based systems—is time-consuming and will probably only be accepted if it offers a clear advantage for the user. This advantage will depend on the context of human-robot interaction scenarios and is more obvious for patient users than for healthy users.

Priors on the expected target objects can easily be integrated into our approach, so that the selection can be viewed as a grounding of the robots' hypotheses of the user's intent. If we would require a minimal threshold for the predicted target probability of decoded object $j*$, the user would also have the option to not attend any object. In an assistive scenario, a robot could thus probe the user to infer if assistance is required at all.

TABLE I

PERFORMANCE OF INDIVIDUAL SUBJECTS USING THE IVAL AND TCOV DECODING METHODS

	5-fold CV target AUC		Online (9 training trials) target AUC		selection accuracy	
	IVAL	TCOV	IVAL	TCOV	IVAL	TCOV
subject						
S1	0.77	0.81	0.68	0.72	0.72	0.9
S2	0.71	0.75	0.63	0.68	0.72	0.82
S3	0.79	0.84	0.66	0.78	0.79	1
S4	0.79	0.9	0.67	0.82	0.79	0.92
S5	0.86	0.89	0.76	0.83	0.92	1
S6	0.77	0.83	0.73	0.78	0.85	0.95
mean ± std	0.78 ± 0.05	0.84 ± 0.05	0.69 ± 0.05	0.77 ± 0.06	0.8 ± 0.08	0.93 ± 0.07

Considering the duration of an object selection, it should be noted that our protocol parameters (choice of training trials and repetitions) has been rather conservative and depending on the desired performance the corresponding numbers could be reduced (c.f., Fig. 4). While we chose a fixed number of repetitions per subject and trial for meaningful comparison of results, an adaptive choice of repetitions in each trial based on the confidence in the predicted target (dynamic stopping) would substantially speed up the selection process [17]. Additionally, detection of error-related potentials after the robot communicates the predicted target object to the user [18] could be exploited to correct potentially erroneous decisions.

VI. CONCLUSION

We present a novel approach to interface-free object selection based on brain signals. The robot elicits event-related responses in the user's brain signal by highlighting objects in the scene using a laser pointer. This novel interaction paradigm does not require user training and avoids unnecessary context switches. Classifying the visual evoked responses from the ongoing brain signals using covariance-based features, a correct target identification of objects was achieved in 93 % of all cases in our online experiments. For future work it would be interesting to include an adaptive choice of object highlighting based on response discriminability as well as reducing the need for user calibration.

REFERENCES

[1] M. A. Goodrich and A. C. Schultz, "Human-robot Interaction: A Survey," *Found. Trends Hum.-Comput. Interact.*, vol. 1, no. 3, pp. 203–275, Jan. 2007.

[2] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K.-R. Müller, "Single-trial analysis and classification of ERP components — A tutorial," *NeuroImage*, vol. 56, no. 2, pp. 814–825, May 2011.

[3] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Classification of covariance matrices using a Riemannian-based kernel for BCI applications," *Neurocomputing*, vol. 112, pp. 172–178, July 2013.

[4] A. Barachant and M. Congedo, "A Plug&Play P300 BCI Using Information Geometry," *arXiv:1409.0107 [cs, stat]*, Aug. 2014.

[5] D. Hübner, T. Verhoeven, K. Schmid, K.-R. Müller, M. Tangermann, and P.-J. Kindermans, "Learning from label proportions in brain-computer interfaces: Online unsupervised learning with guarantees," *PLOS ONE*, vol. 12, no. 4, p. e0175856, Apr. 2017.

[6] M. Schreuder, A. Riccio, M. Risetti, S. Dähne, A. Ramsay, J. Williamson, D. Mattia, and M. Tangermann, "User-centered design in brain–computer interfaces—A case study," *Artificial Intelligence in Medicine*, vol. 59, no. 2, pp. 71–80, Oct. 2013.

[7] M. Tangermann, M. Krauledat, K. Grzeska, M. Sagebaum, B. Blankertz, C. Vidaurre, and K.-R. Müller, "Playing Pinball with non-invasive BCI," in *Advances in Neural Information Processing Systems 21*, D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, Eds. Curran Associates, Inc., 2009, pp. 1641–1648.

[8] K. Muelling, A. Venkatraman, J.-S. Valois, J. E. Downey, J. Weiss, S. Javdani, M. Hebert, A. B. Schwartz, J. L. Collinger, and J. A. Bagnell, "Autonomy infused teleoperation with application to brain computer interface controlled manipulation," *Auton Robot*, pp. 1–22, Feb. 2017.

[9] I. Iturrate, J. Antelis, A. Kubler, and J. Minguez, "A Noninvasive Brain-Actuated Wheelchair Based on a P300 Neurophysiological Protocol and Automated Navigation," *IEEE Transactions on Robotics*, vol. 25, no. 3, pp. 614–627, June 2009.

[10] R. Ying, J. Weisz, and P. K. Allen, "Grasping with Your Brain: A Brain-Computer Interface for Fast Grasp Selection," in *Robotics Research*, ser. Springer Proceedings in Advanced Robotics. Springer, Cham, 2018, pp. 325–340.

[11] A. Tanwani, J. del R Millan, and A. Billard, "Rewards-driven control of robot arm by decoding EEG signals," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug. 2014, pp. 1658–1661.

[12] F. Burget, L. D. J. Fiederer, D. Kuhner, M. Völker, J. Aldinger, R. T. Schirrmeister, C. Do, J. Boedecker, B. Nebel, T. Ball, and W. Burgard, "Acting thoughts: Towards a mobile robotic service assistant for users with limited communication skills," in *2017 European Conference on Mobile Robots (ECMR)*, Sept. 2017, pp. 1–6.

[13] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Multiclass Brain-Computer Interface Classification by Riemannian Geometry," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 920–928, Apr. 2012.

[14] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert, "xDAWN Algorithm to Enhance Evoked Potentials: Application to Brain-Computer Interface," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 8, pp. 2035–2043, Aug. 2009.

[15] F. Yger, M. Berar, and F. Lotte, "Riemannian Approaches in Brain-Computer Interfaces: A Review," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 10, pp. 1753–1762, Oct. 2017.

[16] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. Hämäläinen, "MEG and EEG data analysis with MNE-Python," *Front. Neurosci.*, vol. 7, 2013.

[17] M. Schreuder, J. Höhne, B. Blankertz, S. Haufe, T. Dickhaus, and Michael Tangermann, "Optimizing event-related potential based brain–computer interfaces: A systematic evaluation of dynamic stopping methods," *J. Neural Eng.*, vol. 10, no. 3, p. 036025, 2013.

[18] A. F. Salazar-Gomez, J. DelPreto, S. Gil, F. H. Guenther, and D. Rus, "Correcting robot mistakes in real time using EEG signals," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, May 2017, pp. 6570–6577.