Zero-calibration BMIs for sequential tasks using error-related potentials

Jonathan Grizou, Iñaki Iturrate, Luis Montesano, Manuel Lopes, Pierre-Yves Oudeyer

Zero-calibration BMIs for sequential tasks using error-related potentials

Jonathan Grizou*, Iñaki Iturrate†, Luis Montesano†, Manuel Lopes*, Pierre-Yves Oudeyer*

* INRIA Bordeaux Sud-Ouest, France

{jonathan.grizou, manuel.lopes, pierre-yves.oudeyer}@inria.fr

†Departamento de Informática e Ingeniería de Sistemas (DIIS), Universidad de Zaragoza, Spain

{iturrate, montesano}@unizar.es

Abstract—Do we need to explicitly calibrate Brain Machine Interfaces (BMIs)? Can we start controlling a device without telling this device how to interpret brain signals? Can we learn how to communicate with a human user through practical interaction? It sounds like an ill posed problem, how can we control a device if such device does not know what our signals mean? This paper argues and present empirical results showing that, under specific but realistic conditions, this problem can be solved. We show that a signal decoder can be learnt automatically and online by the system under the assumption that both, human and machine, share the same a priori on the possible signals' meanings and the possible tasks the user may want the device to achieve. We present results from online experiments on a Brain Computer Interface (BCI) and a Human Robot Interaction (HRI) scenario.

I. MOTIVATION

EEG-based brain-machine interfaces (BMIs) (see [1] for a review) provide a communication channel between humans and machines using only brain activity. Since the first demonstration of brain-controlled devices, research on BMIs has emerged as one of the most growing research whose ultimate goal is to endow people with severe motor disabilities with communication and control capacities. Yet, the practical promise of this technology remains unfilled with BMIs remaining confined to the laboratory or limited to clinical studies or home demonstrations that require close technical oversight [2], [3].

Among the existent signals used to develop a BMI, recent works have shown that it is possible to decode information related to human error processing, namely the error-related potentials [4] appearing for instance when the device action does not match the user's expectations. This potential has been used mainly to improve the BMI decoder [5] and, more interestingly, as feedback information to solve sequential tasks [6], [7]. In fact, they can be interpreted as instructions given by a teacher to a learner under the implicit assumption that the two agents share a mutual understanding of feedback meaning (e.g. a decoder that translates raw signals into feedback such as speech into words).

In practice, BMI solves the meaning problem using an open-loop calibration procedure to train a decoder in a supervised manner. This calibration phase hinders the deployment of out-of-the-lab applications [1], due to the need of a specific calibration for each task and session. However, this phase is required mainly due to the non-stationary nature of the EEG [8]; the large intra- and inter-subject variability [9], and variations induced by the task [10].

It is worth noting that this problem can be generalized to many kind of human-machine interaction (HMI) scenarios. In modern human-robot collaboration, there is a need for intuitive interfaces that allow non-technical users to teach robots. A major obstacle is that of requiring a pre-defined set of instruction signals. As a result, in the human to robot interaction (HRI) community, an important part of the work consists of building classifiers to translate human communicative signals (speech, gestures, facial expression) to symbolic meanings understandable by the robot. Such procedure requires a costly offline gathering of signals. A machine able to automatically understand such symbols could improve the usability and ease of use of such interactive system and even make use of involuntary signals, e.g. prosody, to better exploit the information provided by the user.

While research on robot learning from human interaction has flourished in the last ten years [11], most work has focused on how to extract statistical *task models* from human teachers following a fixed pre-defined teaching protocol. Thus, a usual assumption is that the learner and the teacher share a mutual understanding of the meaning of each others' signals. The question of how a robot can learn to interpret personalized and potentially noisy teaching signals, i.e. learn *instruction models*, has been much less explored. In a preliminary work [12], we presented a computational approach addressing this problem by considering a finite space of symbolic teaching signals in simulation while bootstrapping the system with known symbols. Later [13], we released the need for bootstrapping and allow the teacher to use signals that can be represented as fixed length feature vectors, which is better suited for HRI scenarios.

Interestingly, one similar approache have been developed in the BMI community. For non-invasive P300 signals, Kindermans et al. proposed a method to learn from scratch in closed loop a decoder by exploiting multiple stimulations and prior information [14], [15]. However, the approach needed for a warm-up period.

In this paper we present a different approach to calibration procedures. We argue that, for some problems, a calibration procedure is not explicitly needed. The system could learn, by interacting with the user, both what to do and how to map brain signals into meaningful instructions. The innovation of this work is not about new machine learning development

but resides in a new way to use them. Our method combines and exploits two sources of information: task constraints and spacial organization of instruction signals in the feature space. We report results from a BCI online experiments and a HRI scenario. The results show that the proposed method is able to learn good instruction models while solving the task efficiently without any prior calibration procedure.

II. PROBLEM

This section describes the problem of executing a task when the instruction signal's meaning is unknown or uncertain. For practical reasons, we will only refer to feedback driven BMI scenarios for the description of the problem and method. Nonetheless, the idea presented here holds for other HRI scenarios by replacing brain signals with other feedback instructions (e.g., speech). BMI control based on feedback signals (illustrated in figure 1) differs from classical brain-machine interfaces in the sense that the user does not actively deliver commands to the device, but only delivers feedback about actions performed by the device. In this setting, the device needs to actively execute actions to solve the tasks and to be able to learn an intelligent behavior from the feedback. This idea can be seen as a shared-control strategy [16], where both the user and the device help each other to solve a task.
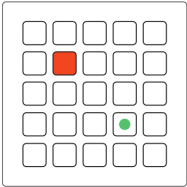

Fig. 1: In this BMI scenario, the user is watching the agent moving on the screen and assess the agent actions with respect to its own objective illustrated by the red state.

Essentially, this BMI control follows an iterative sequential process where, in a particular state s, the device performs an action a and the user assesses the action using brain signal $e \in \mathbf{R}^n$. These assessments generate error-related potentials, i.e. signals elicited in the brain when the outcome observed by the user differs from the expected one [4]. Thus, collected data are in the form $\{(e_i, s_i, a_i), \ i = 1, \ldots, N\}$, i.e. a sequence of states, actions and teaching signals triplets, with N the number of steps. In a common scenario, the system has been fed with a classifier, parameterized by θ, that translates brain signals e into symbolic instructions z that belong to one of two classes (correct or incorrect assessment), $z \in \{c, w\}$. The model parameters θ are in practice learnt using a calibration procedure. The device then learns from this symbolic feedback.

This control can be exemplified for a reaching task (Fig. 1), where the user wants to reach a target position unknown by the system. The device performs several discrete actions (e.g. moving left or right), and learns from the feedback given by the user. After several steps, the device knows which is the desired target and how to reach it. However, the control can become intractable as the task complexity increases. Furthermore, only binary feedback is available and there is a large percentage of misdetected assessments. Given a set of possible tasks $\Xi = \{\xi_1, \ldots, \xi_T\}$, with T the number of task hypothesis, it is possible to speed up the inference by precomputing the agent's optimal behavior $\pi_\xi = p(a|s, \xi)$ for each task and using the feedback signal as a likelihood for the task. For instance, in the previous example, the possible tasks are given by the number of targets. This way, an error (negative feedback) after a particular action will decrease the posterior probability of those targets whose optimal policy agrees with the action.

In this work, we address the problem of removing the need for calibration. Therefore we do not have access to the negative or positive nature $z \in \{c, w\}$ of the brain signals e beforehand. We propose an algorithm that simultaneously calibrates the feedback decoder and executes in closed loop a sequential task only known by the user $\hat{\xi} \in \Xi$. Our method combines and exploits two sources of information: task constraints, namely optimal policies π_ξ, and spatial organization of brain signals in the feature space. The underlying assumptions are:

1) a finite number of possible task hypothesis, i.e. ξ_l for $l \in 1, \ldots, \xi_T$, can be defined 2) the inputs signals have some hidden labels z corresponding to their meaning 3) the set of possible meanings is finite, e.g. $z \in \{c, w\}$ 4) given the ground truth labels $\hat{z}$ of the signals, a classifier of sufficient accuracy could be trained to control the device.

These assumptions may look constraining but are actually common ones in most current BMI scenarios. For example, consider the case of a robotic arm assistant helping to grasp objects on a table. Such robotic assistant could be controlled by a user assessing the robot actions. For instance, the robot could start reaching for an object having the user validating or not the decision of the robot. In such scenario, the usual method would be to start a calibration procedure to map ERP brain signals into symbolic feedback instructions for the robot (correct and incorrect). Once enough data are collected, a classifier would be trained and we could start assessing the robot's actions using brain ERP. This simple scenario, which follows a calibration procedure, already includes all the assumption we defined earlier. We have 1) a finite set of hypothesis represented by the finite number of object on the table 2) a user that is told to assess the robot's actions 3) a finite set of possible meanings (correct and incorrect) 4) signals that can be classified as the calibration procedure was able to generate a usable classifier.

III. METHOD

This section describes our proposed solution to the previously defined problem of executing a task when the instruction meaning is unknown or uncertain and under the given assumption. The main idea is depicted in Fig. 2 for a toy 1D example. The user wants the device to reach the right-most state. However, neither the target $\hat{\xi}$ nor the true feedback labels $\hat{z}$ are known. The feedback signals e are generated as a response to the execution of an action a in state s according to the true unknown task $\hat{\xi}$ the user wants to solve. The key point is that these signals are generated from an underlying model that for binary signals has two different classes. Given sufficient feedback signals, it is possible to build the underlying distributions for each possible target. Only the right task will provide the right meanings (or labels) to each of the feedback signals (Fig. 2 Left), while the other tasks will gradually mix

both classes as the task gradually differs more from the original task (Fig. 2 Middle-Right), up to the point of almost mirroring the labels when the target is mirrored. In the remainder of this section we show how this property can be exploited to estimate the task and the model generating the feedback signals.
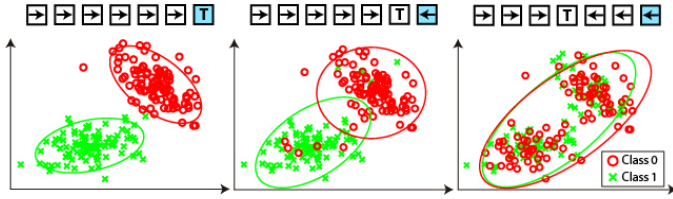

Fig. 2: Task-dependent labels for a 1-D grid world. For the represented example, the arrows indicate for each state what action should elicit a positive feedback to reach the target marked with T (i.e., the optimal policies). 2D Gaussian distributions of binary feedback signals for three possible targets are shown below. While for the correct target the distributions shows a large separability (Left), the overlaps increases as the believed target moves away from the real one (Middle, Right).

Following the literature [17], we will model the EEG signals using independent multivariate normal distributions for each class, $\mathcal{N}(\mu_c, \Sigma_c), \mathcal{N}(\mu_w, \Sigma_w)$. Here, the model parameters θ account for $\{\mu_c, \Sigma_c, \mu_w, \Sigma_w\}$.

Regarding the tasks, the system has access to a set of task hypotheses Ξ which includes the task $\hat{\xi}$ the user wants to solve[1]. We do not make any particular assumption on how the task is represented given that for each particular task ξ we are able to compute a policy π_ξ which represents the probability of choosing a given action a in state s, $\pi_\xi(s, a) = p(a|s, \xi)$. These policies, conditioned on the task, provide meanings to the signals of an action-state pair (e.g. in a reaching task, progressing towards the goal will generate correct answers while moving apart from it will generate wrong ones). We define Z the function that, given a state s, an action a, and a task ξ return the probability of the user intended meaning z, i.e. $Z(s, a, \xi) = p(z|s, a, \xi)$. With $\sum_{f=c,w} p(z = f|s, a, \xi) = 1$. Instead of a binary meaning estimate, we add a noise term to cope with those situations were the user assessment may be wrong. For example, the probability that the user, having task ξ in mind, provides a signal of meaning correct c if the device execute action a in state s is:

$$p(z = c|s, a, \xi) = \begin{cases} 1 - \alpha & if\ a = \text{argmax}_a\ \pi_\xi(s, a) \\ \alpha & \text{otherwise} \end{cases}$$

with α modeling error rate of the user. In our case, only two signal meanings are possible, i.e. correct (c) and incorrect (w), therefore: $p(z = w|s, a, \xi) = 1 - p(z = c|s, a, \xi)$

Following the discussion of Fig. 2, a sensible option to estimate the task $\hat{\xi}$ is to measure the coherence of the signal model θ_ξ computed using the virtual meanings, given by Z, provided by the target policy. In other words, the best (ξ, θ_ξ) pair would provide the lowest predictive error (perr) on the observed signals $p(e|s, a, \xi, \theta)$. One possible way of solving this problem is to maximize the expected predictive

[1]If this is not the case, the system will find the most suitable task.

classification rate:

$$\hat{\xi}, \hat{\theta} \quad = \quad argmax_{\xi, \theta}\ E_e\left(\delta(Z(s, a, \xi), Y(e, \theta_\xi))\right) \quad (1)$$

where $\delta()$ being an indicator function. And $Y(e, \theta_\xi)$ is the predicted label z for signal e under the model parameterized by θ_ξ, i.e. $Y(e, \theta_\xi) = p(z|e, \theta_\xi)$. With $\sum_{f=c,w} p(z = f|e, \theta_\xi) = 1$. In practice, it is just the probability of the meaning under the Gaussian model provided by θ_ξ. For example, the probability that signal e is of meaning correct (c) under θ can be expressed as:

$$p(z = c|e, \theta) = \frac{p(e|z = c, \theta)p(z = c)}{\sum_{k=c,w} p(e|z = k, \theta)p(z = k)}$$
$$= \frac{\mathcal{N}(e|\mu_c, \Sigma_c)p(z = c)}{\sum_{k=c,w} \mathcal{N}(e|\mu_k, \Sigma_k)p(z = k)} \quad (2)$$

In our case, only two signal's meaning are possible, i.e. correct (c) and incorrect (w), therefore: $p(z = w|e, \theta) = 1 - p(z = c|e, \theta)$.

The expected predictive error can be explicitly written dependent on the task and decoder model:

$$E_e\left(\delta(Z(s, a, \xi), Y(e, \theta))\right) =$$
$$\sum_{f=c,w} p(z = f|s, a, \xi)p(z = f|e, \theta) \quad (3)$$

Note that the optimization process has been factored using the fact that given a task ξ, the estimation of θ under the Gaussian model is trivial. It basically requires to compute the maximum-likelihood estimate θ_ξ^{ML} for each task ξ.

Concretely, given a set of task hypothesis Ξ of size T, we can assign, for each hypothesis, probabilistic labels to the signals received from the user (Z). This provides one dataset of signals with T sets of labels. For each task hypothesis and given its associated hypothetic label set, we can now compute the maximum-likelihood model θ^{ML}. By comparing the fitted model prediction (Y) with the initially assigned labels (Z), we can compute a score, here the expected predictive error, that account for the coherence of the spacial organization of brain signals in the feature space with the associated hypothetic labels. The idea is that only the right task will provide the right meanings (or labels) to each brain signals, while the other tasks will gradually mix both classes (Fig. 2). The correct task should therefore have a lower expected predictive error.

IV. RESULTS

In this section we present online results from a BCI scenario as well as a pick and place HRI scenario to illustrate the wider potential application of our approach. For the remaining of this section, we will consider the error rate of the user α equals to 0.1.

A. BCI Control Task

1) Control task: As illustrated in figure 1, we consider a 5x5 grid world, where an agent can perform five different discrete actions: move up, down, left, right, or a target-reached action. The user goal is to teach the agent to reach one, yet unknown to it, of the 25 discrete positions which represent the set of possible tasks. We thus consider that the agent has access

to 25 different task hypothesis (one with goal location at each of the cells). We use *Markov Decision Processes* (MDP) to represent the problem [18]. From a given task ξ, represented as a reward function, we can compute the corresponding policy π_ξ using, for instance, Value Iteration [18].

2) EEG-based feedback signals: EEG signals were recorded with a gTec system (2 gUSBamp amplifiers) with 32 electrodes distributed according to the 10/10 international system, with the ground on FPz and the reference on the left earlobe. The EEG signals were digitized with a sampling frequency of 256 Hz, common-average-reference (CAR) filtered and band-pass filtered at $[0.5, 10]$ Hz.

During operation, the role of the users was to assess the agent actions as good or bad, obtaining this way potentials associated to correct or erroneous actions. Previous studies have demonstrated that these signals can be detected online [4] and even be used as binary feedback signals [6]. Following these studies, features were extracted from two fronto-central channels (FCz and Cz) within a time window of $[200, 700]$ ms (being 0 ms the action onset of the agent) and downsampled to 32 Hz. This leaded to a vector of 34 features. This feature vector served as the input for our algorithm

3) Zero-calibration BCI Control with Human Subjects: This experiment will evaluate the main claim of our algorithm, that we can identify the task desired by the user even without an explicit calibration phase and without any knowledge of the brain signals. The experiments were conducted with four subjects (aged between 25 and 28). Each subject performed 5 runs of learning from scratch how to reach a target (chosen randomly).

Figure 3 summarizes the results. The probability of the correct task (averaged across subjects and tasks) is shown in Fig 3a. Figure 3b shows the run by run results. We can conclude that the algorithm is very robust as all the subjects were able to identify the correct task. There are strong variations among subjects, but we note that in previous works the calibration phase used between 300 and 600 examples [6], [19]. Thus, even for the worst subject, it is still possible to start controlling the system without calibration and in less iteration than required by such calibration procedure.

B. HRI Pick and Place scenario

In this section, we illustrate the broad range of possible application for our approach with a small size pick-and-place task with a real robot. This robot is going to be programmed using a natural speech interface whose words have an unknown meaning and are not transformed into symbols via a voice recognizer. The interaction between the robot and the human is a turn taking social behavior, where the robot performs an action and waits for a feedback instruction signal to continue. This allows to synchronize a speech wave with its corresponding pair of state and action.

1) Experimental System: We consider a six d.o.f. robotic arm and gripper that is able to grasp, transport and release cubes in four positions. We used a total of three cubes that can form towers of two cubes. The robot has 4 actions available: *rotate left*, *rotate right*, *grasp cube* and *release cube*. The state space is discrete and defined as the location of each object,
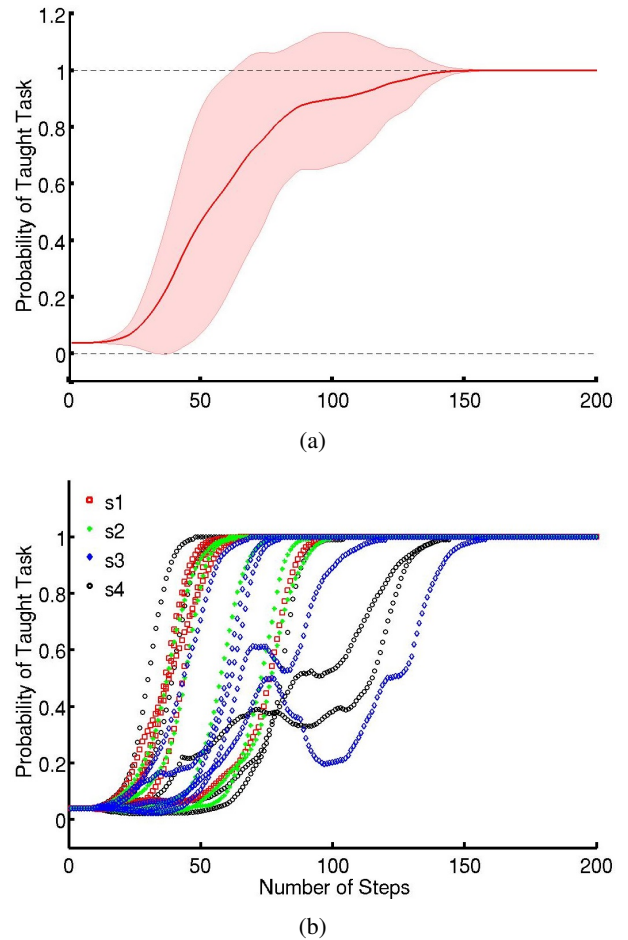

(a)

(b)

Fig. 3: Results from the online BCI experiment for identifying the task: a) Evolution of the probability of the taught task averaged for all subjects; b) Evolution of the probability of the taught task for each subject and run

including being on top of another or in the robot's hand. So for a set of 3 objects we have 624 different states. Figure 4 shows the robot grasping the orange cube.

As for the BCI control task, MDP is used to represent the problem. For this particular representation we assume that the reward function is sparse and so we can generate possible tasks by sampling sparse reward functions. In other words the task is to reach one, yet unknown, of the 624 states of the MDP.

2) Speech processing: We consider speech as the modality for interacting with the robot. After each action we record the teaching word pronounced by the user. This data is mapped into a 20 dimensional feature space using the methodology described next.

A classical method for representing sounds is the *Mel-Frequency Cepstral Coefficients* (MFCC) [20]. It represents a sound as a time sequence MFCC vectors of dimension 12. Comparing sounds is done via *Dynamic Time Warping* (DTW) between two sequences of feature vectors [21]. This distance is a measure of similarity that takes into account possible insertions and deletions in the feature sequence and is adapted
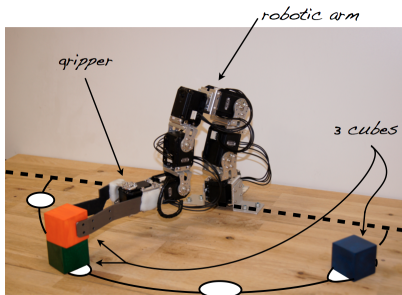
Fig. 4: Robotic System. A six d.o.f robotic arm and gripper learning to performing a pick-and-place task with three cubes.

for sounds comparison of different length. Each recorded vocal signal is represented as its DTW distance to a base of 20 pre-defined spoken words which are not part of words used by the teacher.

By empirical test on recorded speech samples, we estimate that a number of 20 bases words were sufficient and yet a relatively high number of dimensions to deal with a variety of people and speech. This base of 20 words has been randomly selected and is composed of the words: *Error, Acquisition, Difficulties, Semantic, Track, Computer, Explored, Distribution, Century, Reinforcement, Almost, Language, Alone, Kinds, Humans, Axons, Primitives, Vision, Nature, Building*.

It should be made explicit that this is not state of the art speech processing technics but is not the concern of our research. This representation allows to represent spoken words in a relatively low dimensional space with good accuracy.

3) Zero-calibration HRI online pick and place experiment: This brief experiment demonstrates the transferability of our approach to other domains. In addition, it briefly illustrates the ability of our algorithm to reuse acquired knowledge. Once the robot has understood the first task, we can freeze the classifier corresponding to the identified task and start learning a new task faster as this time the signal to meaning mapping is known.
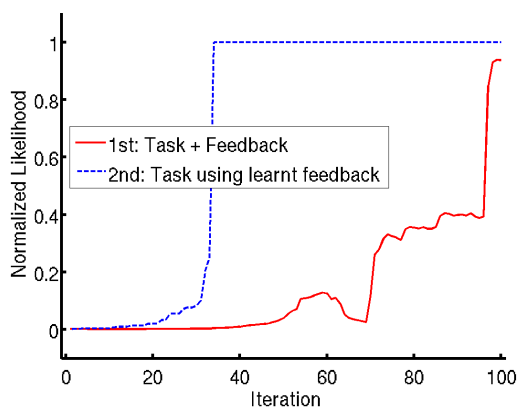

Fig. 5: Evolution of the probability of the taught task. 1) The robot learns a task from unlabeled speech feedback. 2) By freezing the classifier corresponding to the best task estimate, the user teaches the robot a new task faster.

Fig. 5 shows results from one online interactive session with a user using speech to teach the robot what configuration of cube it wanted to the robot to build. In the first run it takes about 100 iterations for the robot to learn the task. Whereas in the second run, when reusing knowledge from the first one, the robot is able to learn a new task faster, in about 30 iterations, meaning that it has well found the two clusters in our feature space as well as the mapping to their corresponding meanings.

V. DISCUSSION

For communication to be successful, the human and the machine need to share some common background which is usually the meaning of the signals received by the device. In practice, such signal to meaning mapping is represented by a specific classifier learnt using a calibration procedure. In this work we have seen that this signal-to-meaning classifier can be learnt automatically and online by the system under the assumption that both, human and machine, share the same a priori on the possible meanings of the signals and the possible task the user may want the device to achieve. We presented a learning algorithm able to associate meaning with unknown signals by reasoning about their relation to previous signals and their relation to the environment itself. The intuition for our method is that the classification of the brain/speech signals is easier when they are interpreted according to the task desired by the user. The method thus relies on finding which pair of classifier-task has the smaller expected prediction error in the signals. We considered the case of brain signals but of particular interest is the possibility to use the same system with other modalities, such as speech. This allows different users to use the system according to their own preferences, skills and limitations. Finally, we showed that, once the system has identified a first task, it can reuse the acquired knowledge about the user instruction signals for learning of a new task faster.

An important challenge for such interactive systems is to deal with non-expert humans. Several studies discuss the different behaviors naive teachers use when instructing robots [22], [23]. An important aspect is that the feedback is frequently ambiguous and deviates from the mathematical interpretation of a reward or a sample from a policy. For instance, in the work of [22] the teachers frequently gave a positive reward for exploratory actions even if the signal was used by the learner as a standard reward. Also, even if we can define an optimal teaching sequence, humans do not necessarily behave according to those strategies [23]. Such aspects were not further considered in this work than by modeling the error-rate of the user.

We believe working without calibration procedure is a novel challenge that can make human to machine interaction more practical to use. Future work will consider how the device can act in order to disambiguate faster the different hypotheses. An important direction is to push this method towards more advanced robotic scenarios by considering, for example, continuous state-action spaces, asynchronous interactions and more complex types of instructions.

REFERENCES

[1] J. Millán *et al.*, "Combining brain–computer interfaces and assistive technologies: state-of-the-art and challenges," *Front Neurosci*, 2010.

[2] E. W. Sellers, T. M. Vaughan, and J. R. Wolpaw, "A brain-computer interface for long-term independent home use," *Amyotrophic lateral sclerosis*, vol. 11, no. 5, pp. 449–455, 2010.

[3] M. A. Lebedev and M. A. Nicolelis, "Brain–machine interfaces: past, present and future," *TRENDS in Neurosciences*, vol. 29, no. 9, pp. 536–546, 2006.

[4] P. Ferrez and J. Millán, "Error-related EEG potentials generated during simulated Brain-Computer interaction," *IEEE Trans Biomed Eng*, 2008.

[5] J. Blumberg, J. Rickert, S. Waldert, A. Schulze-Bonhage, A. Aertsen, and C. Mehring, "Adaptive classification for brain computer interfaces," in *EMBC*, 2007.

[6] R. Chavarriaga and J. Millán, "Learning from EEG error-related potentials in noninvasive brain-computer interfaces," *IEEE Trans Neural Syst Rehabil Eng*, 2010.

[7] I. Iturrate, L. Montesano, and J. Minguez, "Robot reinforcement learning using eeg-based reward signals," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010, pp. 4822–4829.

[8] C. Vidaurre, M. Kawanabe, P. von Bünau, B. Blankertz, and K. Müller, "Toward unsupervised adaptation of LDA for brain-computer interfaces," *IEEE Trans Biomed Eng*, 2011.

[9] J. Polich, "On the relationship between EEG and P300 : individual differences, aging, and ultradian rhythms," *International Journal of Psychophysiology*, 1997.

[10] I. Iturrate, L. Montesano, and J. Minguez, "Task-dependent signal variations in eeg error-related potentials for brain-computer interfaces," *Journal of Neural Engineering*, vol. 10, no. 2, p. 026024, 2013. [Online]. Available: http://stacks.iop.org/1741-2552/10/i=2/a=026024

[11] B. Argall, S. Chernova, and M. Veloso, "A survey of robot learning from demonstration," *Robotics and Autonomous Systems*, 2009.

[12] M. Lopes, T. Cederborg, and P.-Y. Oudeyer, "Simultaneous acquisition of task and feedback models," in *IEEE - International Conference on Development and Learning (ICDL'11)*, 2011.

[13] J. Grizou, M. Lopes, and P.-Y. Oudeyer, "Robot learning simultaneously a task and how to interpret human instructions," *Proceedings of ICDL-Epirob*, 2013.

[14] P.-J. Kindermans, D. Verstraeten, and B. Schrauwen, "A bayesian model for exploiting application constraints to enable unsupervised training of a P300-based BCI." *PloS one*, 2012.

[15] P. Kindermans and H. Verschore, "A P300 BCI for the Masses: Prior Information Enables Instant Unsupervised Spelling," in *NIPS*, 2012, pp. 1–9. [Online]. Available: http://books.nips.cc/papers/files/nips25/NIPS2012_0332.pdf

[16] I. Iturrate, L. Montesano, and J. Minguez, "Shared-control brain-computer interface for a two dimensional reaching task using eeg error-related potentials," in *Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2013.

[17] B. Blankertz, S. Lemm, M. Treder, S. Haufe, and K. Müller, "Single-trial analysis and classification of ERP components: A tutorial," *Neuroimage*, 2010.

[18] R. Sutton and A. Barto, *Reinforcement learning: An introduction*. Cambridge Univ Press, 1998, vol. 28.

[19] I. I, M. L, and M. J, "Single trial recognition of error-related potentials during observation of robot operation," in *EMBC*, 2010.

[20] F. Zheng, G. Zhang, and Z. Song, "Comparison of different implementations of mfcc," *Journal of Computer Science and Technology*, vol. 16, no. 6, pp. 582–589, 2001.

[21] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 26, no. 1, pp. 43–49, 1978.

[22] A. L. Thomaz and C. Breazeal, "Teachable robots: Understanding human teaching behavior to build more effective robot learners," *Artificial Intelligence Journal*, vol. 172, pp. 716–737, 2008.

[23] M. Cakmak and A. Thomaz, "Optimality of human teachers for robot learners," in *Proceedings of the International Conference on Development and Learning (ICDL)*, 2010.