

APPLIED SCIENCES AND ENGINEERING

Utilizing sensory prediction errors for movement intention decoding: A new methodology

Gowrishankar Ganesh,^{1*} Keigo Nakamura,¹ Supat Saetia,² Alejandra Mejia Tobar,² Eiichi Yoshida,¹ Hideyuki Ando,³ Natsue Yoshimura,^{2,4} Yasuharu Koike²

We propose a new methodology for decoding movement intentions of humans. This methodology is motivated by the well-documented ability of the brain to predict sensory outcomes of self-generated and imagined actions using so-called forward models. We propose to subliminally stimulate the sensory modality corresponding to a user's intended movement, and decode a user's movement intention from his electroencephalography (EEG), by decoding for prediction errors—whether the sensory prediction corresponding to a user's intended movement matches the subliminal sensory stimulation we induce. We tested our proposal in a binary wheelchair turning task in which users thought of turning their wheelchair either left or right. We stimulated their vestibular system subliminally, toward either the left or the right direction, using a galvanic vestibular stimulator and show that the decoding for prediction errors from the EEG can radically improve movement intention decoding performance. We observed an 87.2% median single-trial decoding accuracy across tested participants, with zero user training, within 96 ms of the stimulation, and with no additional cognitive load on the users because the stimulation was subliminal.

INTRODUCTION

The big challenge for brain computer interfaces (BCIs) that aim to help humans in motor tasks is the deciphering of a human user's movement intention from his brain activity while minimizing user effort. Although a plethora of methods have been suggested for this in the past two decades (1–3), they are all limited by the effort demanded from the human user—the so-called reactive BCIs decode movement intention by analyzing the observed response in electroencephalography (EEG) to an explicit external stimulus (4, 5). This involves an additional attentional and cognitive load for users who need to attend to a very conspicuous stimulus, in addition to the task they want to control using the BCI (6). Active BCIs on the other hand, which decode movements using the inherent EEG activity before movements (2, 3, 7–10), require extensive user training, either to enable the user to control his or her brain activity (11, 12) or to imagine movements (2, 3). Furthermore, the imagined movements, which are popularly the movements of lateral limbs, are often different from the task they want to control using the BCI, putting again an additional cognitive load on the users. Here, we propose a new (subliminal) stimulus-based reactive BCI methodology that is fundamentally different from all the previous methods and promises to drastically attenuate many of these user-related constraints. We show that the decoding based on this new method can provide high single-trial decoding accuracies across participants, with zero user training, and with no additional cognitive load on the users.

The key difference between all the previous methods and what we propose is in what is decoded. All the previous methods decode what movement a user intends/imagines, either directly (as in the active BCI systems) or indirectly by decoding the stimulus he or she is attending to (like the reactive BCI systems). Here, we propose to not decode what movement a user intends/imagines but to decode whether the movement he or she wants matches the sensory feedback we induce

using a subliminal stimulator. This idea is motivated by the multitude of studies on forward models (the neural circuitry implicated in predicting sensory outcomes of self-generated movements) (13). The sensory prediction errors, between the forward model predictions and the actual sensory signals, are fundamental to our socio-motor abilities. They are known to play crucial roles in the control of self-generated actions (13–15), haptic perception (16), motor learning (17), and probably even interpersonal interactions (18–21) and cognition of self (22). Sensory predictions are known to be present not only during movement but also during the imagination of movement (23, 24), that is, our brain expects sensory signals corresponding to imagined movements. Because of these extensive functionalities, we hypothesized the sensory prediction errors to have a significant signature in the EEG and, consequently, that the decoding for the presence of a prediction error (that is, whether a participant's movement intention matches the sensory feedback he or she feels) to be a good proxy for decoding what movement a participant intends.

We tested this hypothesis in a simulated wheelchair turning task, where participants (sitting on a chair) were asked to imagine the wheelchair turning either right or left. Vestibular feedback is the most dominant sensory modality during turning. Hence, in each trial, we stimulated the participants' vestibular system so as to induce a sensation of either turning left or right, albeit subliminally, using galvanic vestibular stimulation (GVS) (25–27). We show that even a subliminal stimulation induces prediction errors, which we are then able to decode, and, knowing the stimulation direction, consequently decode the direction the participant intends to turn. Figure 1A shows the concept of the proposed technique with regard to the wheelchair experiment.

RESULTS

Twelve participants participated in our experiments. The participants sat on a chair in a noise-insulated room with their eyes closed and were cued with an audio signal by either of two speakers, placed on either side of the participant. They were instructed to “think about wanting to turn toward the speaker and imagine the feeling that their body turning accordingly.” Two seconds after the end of every audio cue, a GVS (that corresponded to either right or left rotatory vestibular sensation) was

Copyright © 2018
The Authors, some
rights reserved;
exclusive licensee
American Association
for the Advancement
of Science. No claim to
original U.S. Government
Works. Distributed
under a Creative
Commons Attribution
NonCommercial
License 4.0 (CC BY-NC).

¹CNRS-AIST (Centre National de la Recherche Scientifique–National Institute of Advanced Industrial Science and Technology) Joint Robotics Laboratory (JRL), UMI3218/RL, Tsukuba Central 1, 1-1-1 Umezono, Tsukuba, Japan. ²Tokyo Institute of Technology, Tokyo, Japan. ³Osaka University, Osaka, Japan. ⁴Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Agency (JST), 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan.
*Corresponding author. Email: gans_gs@hotmail.com

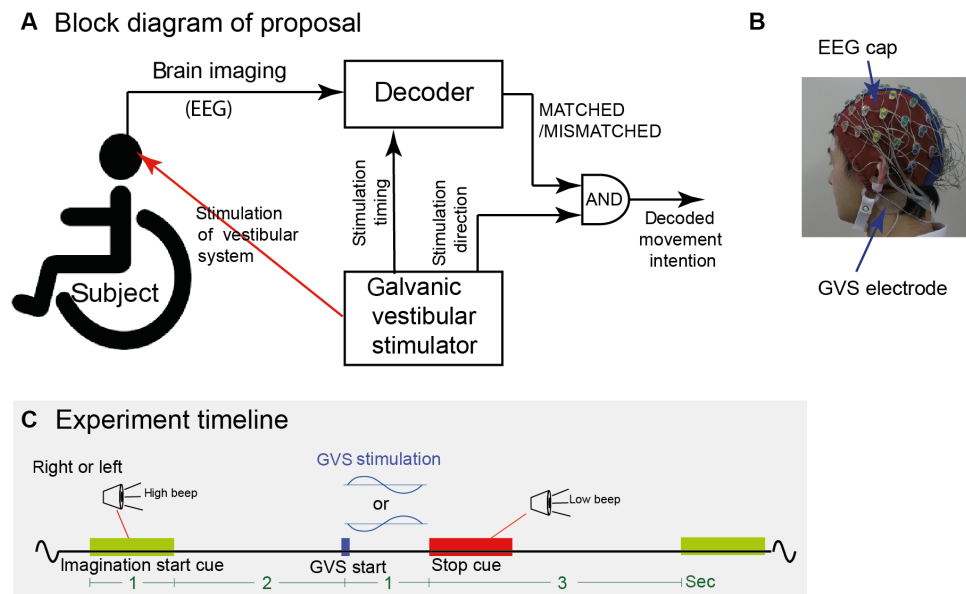


Fig. 1. Proposal and experiment paradigm. (A) We propose to use a sensory stimulator in parallel with EEG and decode whether the stimulation matches the sensory feedback corresponding to the user's motor intention. The presented experiment simulated a wheelchair turning scenario and used a galvanic vestibular stimulator. (B) The participants were affixed with GVS electrodes during the EEG recording. The subliminal GVS induced a sensation of turning either right or left. (C) Experiment timeline: In each trial, using stereo speakers and a "high"-frequency beep, the participants were instructed to imagine turning either left or right while sitting on a rotating chair. A subliminal GVS was applied 2 s after the end of each cue, randomly corresponding to turning either right or left. This was followed by a rest period of 3 s cued by a "low"-frequency beep (stop cue).

provided behind the participant's ears (Fig. 1B, and see experiment timeline in Fig. 1C). The vestibular stimulation was subliminal, such that none of the participants perceived any vestibular sensation (please see Validation that GVS was subliminal section). EEG was recorded from the participants through the experiment. The detailed experimental procedure is provided in the detailed methods.

Intention decoding

First, we collected the EEG data from each trial and aligned it to the intention start (auditory) cue presentation time. The cue aligned data were divided into two classes labeled according to whether the cue was from the right (RHT-CUE) or left (LFT-CUE) speaker, that is, whether the participants thought of turning right or left. We then considered different time periods of EEG data (either 500, 1000, 1500, 2000, or 2500 ms of data from the cue presentation) and tried to decode the participant intention using the absolute magnitude of the EEG signals across these time periods (collected at 512 Hz) as features (we will discuss results of a frequency-based decoding later). Because of its robustness to overfitting, a sparse logistic regression (SLR) (28) decoder was chosen, and trained on part of the data (80% of randomly selected trials) to classify the turning direction the participant thinks of, and tested on a test set (remaining 20% trials). This procedure was repeated 20 times for each participant to achieve a mean decoding performance with each participant.

The red trace in Fig. 2 plots the across-participant decoding performance as box plots. The decoding data were normal across participants in all the time periods ($P > 0.05$, Shapiro-Wilk test), allowing us to perform an analysis of variance (ANOVA). The intention decoding did not change with the time period ($F_{4,24} = 0.85$, $P > 0.50$, one-way repeated-measures ANOVA) and did not improve above chance ($t_{34} = 0.133$, $P = 0.89$, one sample t test on the difference of the data from all participants and time periods, from the chance level of 50%).

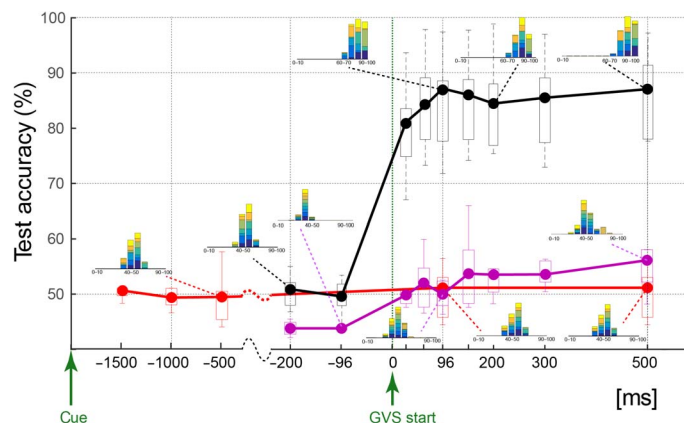


Fig. 2. Decoding performance summary. The across-participant median decoding performance when decoding for the direction in which a participant wants to turn (that is, the cue direction) is shown in red and pink, whereas decoding for a MATCHED/MISMATCHED participant intention and applied GVS is shown in black. The data at each time point represent the decoding performance using data from the time period between a reference point ("cue" for red data and "GVS start" for pink and black data) and that time point. Box plot boundaries represent the 25th and 75th percentile, whereas the whiskers represent the data range across participants. The inset histograms show the participant ensemble decoding performance in the 140 (20 test trials \times 7 participants) test trials, with each participant data shown in a different color.

Prediction error decoding

Next, we aligned the EEG data to the start of the GVS. Crucially, this time, we collected the data into two classes that were labeled as either MATCHED, if the cue and GVS directions corresponded, or MISMATCHED, if the cue and GVS directions did not correspond. Note that both the MATCHED and MISMATCHED classes contained

(equal number of) trials of both the left and right GVS. With these data, we checked the prediction error decoding performance; we trained a decoder to classify whether a trial was MATCHED or MISMATCHED considering different time periods of EEG from before $\{-200, -96\}$ ms and after $\{32, 64, 96, 150, 200, 300, 500\}$ ms of GVS start. Again, in each case, the absolute EEG signal magnitudes across these time periods were used as features for the decoder.

The black trace in Fig. 2 shows the across-participant results of the MATCHED/MISMATCHED decoding. The decoding results from all the time periods were again normal across participants ($P > 0.05$, Shapiro-Wilk test). The decoding shows chance performance before start of GVS ($t_6 = 0.44$, $P > 0.67$), but rises sharply after start of GVS. The performance peaks when using 96 ms of EEG data after GVS, to a median of 87.6% (ranging between 72 and 97.9%) across participants, and remains similar thereafter ($F_{4,24} = 2.22$, $P > 0.09$, one-way repeated-measures ANOVA on the decoding results after 96 ms). We will therefore concentrate on the 96-ms time period for the detailed analysis hereafter.

Comparing intention decoding and prediction error decoding

The decoding values during prediction error (MATCHED/MISMATCHED) decoding were significantly higher than chance using 96 ms of data after GVS start ($t_6 = 10.43$, $P < 4.6 \times 10^{-5}$, one-sample t test; see black data in Fig. 2) and significantly higher than direct intention decoding (red trace) performance at the same time ($t_6 = 12.87$, $P < 1.35 \times 10^{-5}$). Because the intention decoding (red data in Fig. 2) was performed relative to the cue (that was presented between 3 and 2 s before the GVS), it used data across longer time periods (and hence a larger number of features) for training. Although we use a sparse classification algorithm for our decoding, it is possible that the large feature volume led to overfitting and, consequently, lower performance. Therefore, we also attempted to decode the intention (that is, cue) within the same time periods (relative to the GVS start) as used for the prediction error decoding. These data are shown in pink. Again, we observed that the direct intention decoding remained close to chance. A two-way repeated-measures ANOVA across the two decoding strategies (intention decoding and prediction error decoding) and time periods exhibited a significant main effect of decoding strategy ($F_{8,48} = 39.3$, $P < 10^{-5}$), although there was also an effect of time periods ($F_{8,48} = 49.86$, $P < 10^{-5}$) and an interaction ($F_{4,6} = 152.14$, $P < 1.76 \times 10^{-5}$), which are obvious because of the jump in decoding accuracy during prediction error decoding. Post hoc t tests revealed that the prediction error decoding accuracy was greater than intention decoding in every time period after 96 ms of GVS start ($t_6 = 7.85$, $P < 2.2 \times 10^{-4}$; 96-ms period data are plotted in Fig. 3).

Finally, we noted that the intention decoding we have performed until now uses time domain EEG features, whereas recent studies have shown that frequency features can help provide very high event-related desynchronization (ERD)-based decoding of motor imagery (29, 30). Therefore, we also tried to decode the participant intention in our task using frequency features. Considering the periods between the cue start and the GVS, we used a 1-s window (shifted by 100 ms) to evaluate the power spectrum in the range of 0.01 and 30 Hz and used these as features to decode for the participant intention. We observed an across-participant median decoding accuracy of 48.85% (ranging between 46.77 and 52.82%; see orange plot in Fig. 3), which was not different from chance ($t_6 = 0.52$, $P = 0.62$). This result suggests that ERDs were not excited in our task.

Overall, these results support our proposal of using subliminal sensory stimulator with EEG and decoding for the prediction errors. Note that once the prediction errors (MATCHED/MISMATCHED classes) are decoded, inferring the actual participant intention is trivial because the stimulation direction is known precisely. Furthermore, our results show that prediction error decoding can work even in cases where active decoding methods (like ERD) do not provide good results.

Thinking of turning or not

A common question in BCI setups for movement decoding is to know when to perform the decoding, that is, to know when a participant is thinking of a movement. Our preliminary investigation shows that the sensory stimulation can help in this regard. Our experiment included so-called no-intention trials (see Materials and Methods), in which the GVS was applied during rest, when the participants were not asked to think of turning in either direction. We observed that using 96 ms of EEG signals (after GVS start), we could decode with appreciable accuracy (median 74.21% across participants; see cyan plot on Fig. 3) whether a participant was thinking of turning (either right or left) or not (that is, it was a no-intention trial).

Spatiotemporal characteristics of features

We also analyzed which spatiotemporal features were selected by the decoder for the MATCHED/MISMATCHED decoding in the 96-ms time period after GVS start. The decoder we used is designed to minimize the selected features. Therefore, the selected features may not represent the entirety of brain activity related to prediction errors. On the other hand, the decoder-selected features (especially ones common between participants) are robust to false positives, and hence, we

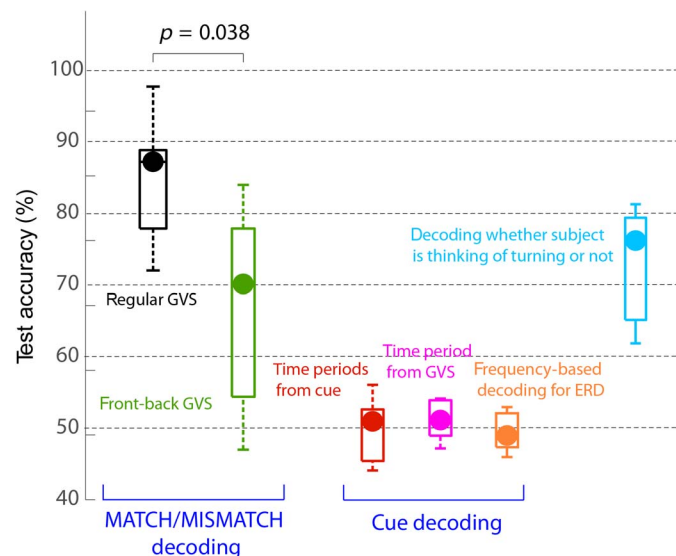


Fig. 3. Summary of decoding performance from various analyses in the 96-ms time period. The MATCHED/MISMATCHED decoding with GVS is shown in black. On applying the GVS in the front-back configuration (green data), the decoding performance decreased. Decoding for cue direction, using either all the data since the cue (red data) or data since the GVS start (pink), did not show performance above chance. Decoding performance was similarly low using frequency features between 0.01 and 30 Hz, typical of ERD-based decoders (orange), indicating that our task (imaging turning right or left) did not initiate significant ERD differences in participants. Finally, the decoding of whether the participant was thinking of turning or not (cyan) was also observed to be significant after GVS.

can probably say with some certainty that the selected channels represent brain activity related to prediction errors.

The mean decoder weight assigned to each channel across the time period and participants is shown in Fig. 4A. A darker color represents a higher weight (normalized by the maximum weight). Nine EEG channels, specifically F7, FPz, AFz, Fz, AF8, F8, CPz, P4, and O2 were selected as features in at least six of the seven participants (see channels highlighted by blue disks in Fig. 4A). Figure 4B shows the temporal period across the participants when these nine channels were selected.

Control to check GVS artifacts in the decoding

Note that we do not remove the GVS artifacts in the channels and are able to perform the MATCHED/MISMATCHED decoding in the presence of the GVS artifacts. However, each of the MATCHED and MISMATCHED classes includes equal trials of both the right and left GVS. Artifacts from the GVS therefore cannot theoretically help the decoding. However, to further check that this is indeed the case, we performed a second experiment with five participants in which the GVS was applied, not across the ears (which is optimal to excite the left or right vestibular stimulations) but across the forehead and the back of the neck of the participants (in the forward and backward directions). We aimed to check two issues with this experiment. First, whether we indeed decode prediction error: In a front-back GVS configuration, neither direction of GVS would properly correspond to the participant's intended turning direction (left or right). This in effect should decrease the difference between the MATCH and MISMATCH classes and, hence, reduce our decoder performance. Second, if the GVS artifacts are crucial to the decoding, then the decoder-selected features are expected to change with the change in the GVS location. On the other hand, if the decoder is GVS-irrelevant (like we claim), then the selected features should still match the features selected in Fig. 4.

We found that this was indeed the case. The second experiment was similar to the main experiment in every aspect of the procedure and data analysis, except for the GVS location. The participants in the experiment still imagined either turning to the left or the right, and for our decoding, we labeled the trials with right cues and forward GVS, and left cues and backward GVS as MATCHED trials. The other trials were labeled as MISMATCHED. Even with the front-back GVS, the MATCHED/MISMATCHED decoding accuracy was still significant (median, 71.7%; ranging between 48.13 and 85.83%) across participants (green data in Fig. 3). This was expected because although we applied the GVS across the forehead and neck, we can expect that it activates the rightward or leftward vestibular sensation as well to some extent. However, crucially, the prediction error decoding was significantly less than the case when the GVS corresponded to the participant intention ($t_{10} =$

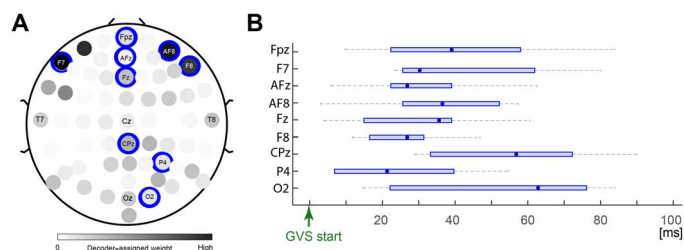


Fig. 4. Spatiotemporal features selected in the 96-ms time period. (A) The weights selected by the decoder, averaged across the channels and time, are plotted in shades of gray. Nine channels (indicated by the blue disks) were selected in six or more participants. (B) The across-participant median (data point) and time range (whiskers) in which these nine channels were selected.

$2.38, P = 0.038$, two-sample t test between the decoding accuracy between the two experiments).

Furthermore, although the decoder performance was lower, the features selected by the decoder in the participants were very similar to those selected in the main experiment—a very strong correlation was observed between the EEG channel weights chosen by the decoder in the front-back GVS configuration and those chosen (Fig. 4A) in the main experiment (Pearson's $R = 0.56, t_{62} = 5.322, P < 10^{-5}$). This result is critical because it not only ratifies the role of these features in prediction error decoding but also, as previously mentioned, shows the independence of the features to the GVS location, suggesting that GVS does not influence the decoding.

Validation that GVS was subliminal

Finally, we performed an additional experiment with 10 participants (7 previous and 3 new participants) to quantify any perceivable sensation due to the GVS (Fig. 5; see Materials and methods for details). The experiment required participants to report the level of any felt “vestibular perturbation,” “muscle twitch,” or “tactile sensation” with a score of 0 to 6 on the Likert scale (0 represents no report) while we applied increasing magnitudes of GVS. The plot in Fig. 5 combines the participant reports during the leftward and rightward GVS such that the solid trace represents the median across 20 data points (10 participants \times 2 GVS directions). The box plot shows the 75% percentile and range (whiskers) of the participant reports. All the participants answered zero for all the three sensations for the level of 0.8 mA (that we use in the experiment). However, some participants did report a “slight” tingling during GVS. This is a previously reported phenomenon that occurs in the presence of dry skin or hair under the electrodes. The tingling was, however, nondirectional (did not change with direction of GVS) and could be attenuated by shaving the hair under the electrodes and wetting the skin with a swab.

Figure 5 reports the scores from when the GVS was applied for 1 s. However, we show that, in fact, less than 100 ms of stimulation is required for the motor intention decoding. Corresponding to this, we also conducted a second validation session in which the 1-Hz GVS was cut off after 100 ms. In this session, none of the participants felt any sensation; the score remained zero for vestibular perturbation, muscle twitch, and tactile sensation; and no tingling was reported even up until

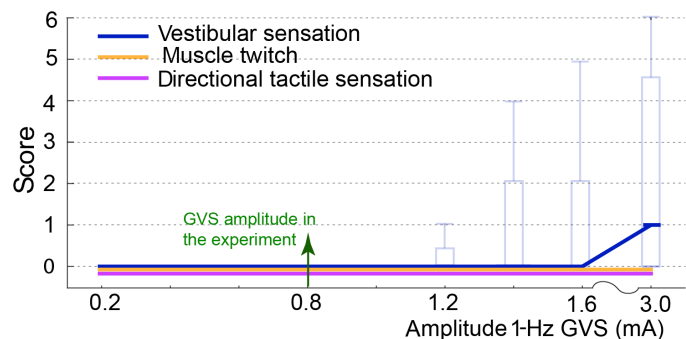


Fig. 5. GVS perception questionnaire. The report of vestibular sensation, muscle twitch, and tactile stimulation in our experiment to validate the GVS was subliminal. An experimenter applied GVS at 1 Hz, of amplitudes increasing from 0.2 to 3 mA (in both the rightward and leftward directions). Participants scored any sensation they experienced, between 0 and 6 (0 representing no report). The box plots show the across-participant median, the 75% percentile, and range (whiskers) of the participant reports. Participant reports were zero for all the sensations at the amplitude of GVS (0.8 mA) used in our experiment.

3 mA (which is almost four times the stimulation applied in our experiment). These results show that the GVS applied in our experiment were subliminal, and the participants did not perceive any significant sensation due to it.

DISCUSSION

Studies over the past two decades have repeatedly shown that sensory predictions by the so-called forward models modulate various aspects of the human behavior (13–22). On the basis of these findings, here, we proposed a new decoding methodology to decode the movement intention of a participant by decoding for prediction errors—the match or mismatch between the sensory prediction corresponding to a user's intended movement and a corresponding sensory stimulation. We hypothesized that the decoding of prediction errors is a good proxy for the decoding of the intention. The significantly better decoding of the prediction error (MATCHED/MISMATCHED data) compared to decoding of cue direction (Figs. 2 and 3) supports our hypothesis. Crucially, the high decoding accuracy was observed over individual trials, without any participant training, within tens of milliseconds of the sensory stimulation, and using subliminal sensory stimulations that the user could not cognitively perceive. This exhibits the promising potential of our proposed method of using sensory stimulation in parallel to brain imaging for movement intention decoding.

There exist several fundamental differences between the prediction error (MATCH/MISMATCH) signals we detect here and the well-established signatures of mismatch negativity (MMN), error-related negativity (ERN), and ERD in EEG. MMN, which was first reported for auditory signals (31) and has also been investigated in visual signals (32), is considered to be a consequence of perceivable sensory expectations induced by previous (often repetitive) sensory experiences. However, the participants in our task were not exposed to any perceivable or repetitive sensory signal, as is typical of MMN studies (33, 34). Consequently, the decoding we performed here is explained better as a mismatch of predicted sensory signals relative to one's motor intention, rather than previous sensory experiences.

ERN, on the other hand, represents a neural response to the commission of an error. Crucially, ERN is activated after an erroneous action by a participant that leads to perceivable sensory stimulations (35, 36). In stark contrast, we decode intended actions before they are performed and using subliminal sensory stimulations. Furthermore, ERN is considered to be generated in the anterior cingulate cortex and usually detected in the mid brain EEG channels Cz and FCz (35). These channels do not correspond to the features selected frequently by our decoder.

Finally, participant intention is popularly detected using ERD (29, 30, 37). ERDs are known to be induced by the imagery of motor actions. ERD-based decoding is, however, not effective for all participants and tasks. In our task as well, ERD-based decoding was not effective (see orange data in Fig. 3), indicating that our task (imagining one's body turning right or left) did not initiate significant ERD differences in participants. On the other hand, the fact that MATCHED/MISMATCHED decoding worked well suggests that the prediction errors are complementary to ERD and may be integrated with ERD-based decoding to improve intention decoding.

Because the features selected by a decoder are prone to false negatives, we cannot deduce much about the spatial characteristics of the underlying brain activations from the electrode features selected by our MATCHED/MISMATCHED decoder (Fig. 4). However, it is still

interesting to note that the frequent selection of the parietal electrodes agrees with the reported involvement of the parietal cortex in sensory estimation during actions (38, 39) and the vestibular processing for motor control in the right interparietal sulcus (40). Any prediction error related to vestibular signals is thus expected to also be represented in these areas. The frontal cortex features are consistent with a recent study that found that the activity in the frontal midbrain channels elicits activations (within 100 ms) representing the absence of associated sensory stimuli after a motor action (41).

Notably, whereas prediction error decoding gives very good decoding performance, what we present here is a prototype system that can still be improved in several aspects. First, the GVS that we provide is not what is optimal for our task. We ask our participants to imagine their body turning about the longitudinal axis, but because of technological limitations, the GVS that we are able to provide predominantly (but not exclusively) induces a vestibular sensation of rotation about the sagittal axis. Because of the centrifugal forces on the head, turning one's body about does lead to a sagittal vestibular sensation as well, and therefore, the GVS we provide is not unrealistic (as also suggested by our results). However, it also shows that there is scope for improvement. Second, for the current proof-of-concept experiment, we chose a very simple sinus temporal GVS profile and used a very low GVS amplitude (see Materials and methods for details) to ensure it is subliminal for all participants. The shape and amplitude can be optimized and customized to individual participants to improve the decoding. Third, here, we used the SLR (28) algorithm because it is parameter-free and robust to overfitting. SLR, however, is known to over-prune features (42), and hence, performance may be improved using other decoding algorithms. We did try the decoding using support vector machines [Matlab Support Vector Machine (SVM) toolbox] and iterative SLR (iSLR) (42), both of which though gave similar decoding performances. Finally, because our proposed method requires a sensory stimulation when the participant is thinking about a motor action (turning his or her body in our current experiment), an obvious question in this scenario is how to know when a participant is thinking about an action. This in fact is a common question in all BCI setups. Our preliminary investigation shows that our proposed methodology can help both detect whether a participant wants to make a movement (blue plot in Fig. 3) and detect which direction he or she wants to turn (black plot in Fig. 3).

In conclusion, we proposed a new methodology for motor intention decoding, that of using a subliminal sensory stimulator with EEG, using a random stimulation and decoding the prediction error between an intended action and the subliminal sensory stimulation. We present this procedure in a two-class wheelchair turning scenario. In this scenario, we used galvanic vestibular stimulation, because vestibular feedback is the dominant feedback during body turning. Other possible applications of this methodology may be for the control of a prosthetic limb by an amputee, the control of an additional robotics limb after stroke or paralysis, or the control of exoskeletons and functional augmentation suites by the elderly. For each of these applications, one would need to stimulate the proprioceptive or tactile sense that corresponds to the action, such as by using a peripheral nerve stimulator (43) or emerging technologies of direct brain stimulation (44). An important application of EEG-based BCI is for communication with locked-in patients (1). Forward models are known to operate during verbal communications. Therefore, in the future, when a suitable technology to stimulate the auditory cortex for specific words or sounds is developed, we would like to imagine that our methodology can be used for speech production, that is, enable locked-in patients to speak. However, for now,

our proposal promises to improve motor imagery–based techniques [for example, Höhne *et al.* (45)] used to communicate with locked-in patients. Overall, the key contribution of our proposed method is the high decoding accuracies within 96 ms of stimulation, without any participant training and without any cognitive load on the participant. These features show its promise for online decoding and for improving current BCI motor augmentation systems by integrating the proposed method with available decoding methods.

MATERIALS AND METHODS

Experiment

Participants

Twelve participants (2 females) aged between 20 and 51 participated in our experiments. Seven participants participated in the main experiment, whereas five participated in a control experiment with suboptimal sensory stimulation. Participants gave written informed consent before participating, and the study was conducted with ethical approval of the ethics committee at the Tokyo Institute of Technology in Tokyo, Japan.

Materials

The experiments were conducted in a noise-insulated room. We used a commercial EEG recording system (ActiveTwo amplifier system; 64 active sensors, Biosemi) for recording the brain activity at 512 Hz. A custom-made wireless galvanic vestibular stimulator (25–27) was used in parallel to the EEG recording. Participants were presented with audio cues using stereo speakers to instruct what they should think/imagine (please see Data collection section). The experiment cue presentation and synchronization between the cue presentation, the GVS system, and the EEG recordings were achieved using a Matlab-based program.

Protocol

Our experiment lasted around 2 hours in total, including time for GVS electrode placement, EEG electrode setup, participant instruction and familiarization, data collection, and subsequent cleanup.

Electrode fixation and familiarization. The experiment started with the fixation of the GVS electrode pads behind the participant's ears and above the temporomandibular joint. With each participant, the GVS was applied a few times to confirm that the stimulation was not perceived by the participant and he or she did not feel any itching/tingling (present sometimes if the hair below the electrode pad is long). In cases where the participant felt itchy, the hair below the pads was shaved, and the electrode pads were refixed. None of our participants reported any perceived vestibular stimulation. The EEG cap and electrodes were fixed next. The participants were then asked to sit on a revolving office chair inside a noise-insulated room. They were instructed on the experiment cues (explained in the next section).

Data collection

EEG data were recorded from each participant over six sessions, each lasting about 6 min. Each session consisted of 50 trials, each with a timeline, as shown in Fig. 1C. After an initial rest of 10 s, each trial started with a high-pitched (around 1000 Hz) audio cue (which lasted 1 s) that was presented on either of the two speakers placed on the left or right side of the participant. The participants were instructed to “think about wanting to turn toward the speaker and imagine the feeling corresponding to your body turning accordingly.” This instruction was repeated before every session. To help the participants understand this instruction, a short demonstration was performed in which the participants were asked to close their eyes and feel their body revolving as the experimenter turned the revolving chair (with the participant in it) left and right alternatively. This procedure was performed for about 20 s

before starting the first session, and again in between sessions if the participants requested it when given the option.

After 3 s of the high-pitched cue, a low-pitched (around 200 Hz) beep, presented on both speakers, instructed the participant to “stop thinking and relax.” The rest lasted for 3 s. GVS was applied after 2 s of the high-beep cue. We applied a sinus GVS signal with anode either below the right ear (right GVS) or below the left ear (left GVS). The GVS sinus amplitude and frequency were set to low values of 0.8 mA and 1 Hz, respectively, to avoid any possible perception. The direction of the cue (instructed direction of thinking) and GVS direction through the session were presented in a fixed pseudorandom order (unknown to participant) such that, in each session, the participants experienced the following: (i) 10 randomly presented CRGR trials, in which they were presented with an audio cue to their right and thought about turning right, and we applied a subliminal GVS corresponding to a rightward turning vestibular sensation; (ii) 10 randomly presented CLGL trials, in which they were presented with an audio cue to their left and thought about turning left, and we applied a subliminal GVS corresponding to a leftward turning vestibular sensation; (iii) 10 randomly presented CRGL trials, in which they were presented with an audio cue to their right and thought about turning right, but we applied a subliminal GVS corresponding to a leftward turning vestibular sensation; (iv) 10 randomly presented CLGR trials, in which they were presented with an audio cue to their left and thought about turning left, but we applied a subliminal GVS corresponding to a rightward turning vestibular sensation; and (v) 10 randomly presented no-intention trials, in which they were not presented with an audio cue and, hence, did not think about turning either direction while we applied a subliminal GVS corresponding to either a rightward or leftward turning vestibular sensation (five trials in each direction). These trials were used to evaluate if the GVS could help us decode whether the participant is thinking about turning (either left or right). We labeled the CRGR and CLGL trials as “MATCHED” trials, and the CRGL and CLGR trials as “MISMATCHED” trials.

The cue and GVS timings were optimized with preliminary experiments to minimize experiment time while allowing sufficient time for the participant to start thinking after each cue. The time between the cue and GVS was purposely kept long (>2 s) to enable us to compare the performance of decoding the cue to the decoding of the prediction error. The GVS profile was also tuned with preliminary experiments. Note that the 1-s GVS profile is still not optimal and, as will be shown in the data later, 100 ms of GVS is enough to get appreciable decoding accuracies.

Following the experiment, the participants were asked again if they perceived any vestibular stimulation during the experiment. The EEG electrodes were then removed from the cap, the EEG cap and GVS electrodes were removed, and the participants were allowed to clean up.

Additional experiment to validate GVS was subliminal

To validate that the GVS applied to the participant were indeed subliminal, we conducted an additional perception experiment [similar to that of Oppenlander *et al.* (46)] with 10 participants (7 of our previous participants and 3 additional new participants; we could not manage to recall five of our previous participants). The GVS electrodes were fixed behind the participant's ears, similar to our main experiment. The participants were then asked to close their eyes while an experimenter applied GVS at 1 Hz and for 1 s (same as our experiment). The amplitude of GVS was increased from 0.2 to 3 mA as {0.2, –0.2; 0.4, –0.4; 0.6, –0.6; 0.8, –0.8; 1.0, –1.0; 1.2, –1.2; 1.4, –1.4; 1.6, –1.6; 3.0, –3.0}. The

order of the positive (rightward) or negative (leftward) GVS of the same amplitude was chosen randomly, and the time gap between the two stimulations was randomized so that the participants could not predict when a stimulation was applied. The participants were asked to interject the experimenter whenever they felt any sensation of vestibular perturbation, muscle twitch, or/and “direction tactile sensation,” and report the level of the sensation on a Likert scale of 1 to 6 (0 defining no report). The box plot of the reported scores for the three sensations is shown in Fig. 5. In the second session of the experiment, the same protocol was followed, except that the 1-Hz GVS was cut off after 100 ms.

EEG data processing

After the preliminary analysis, we chose temporal features from raw EEG data from each of the 64 channels as features. We checked the decoding performances with three decoding algorithms: a recent popular SLR algorithm (28), SVM (Matlab toolbox), and the iSLR algorithm (42). The results were found to be equally good with the three algorithms. We will discuss the SLR results in detail here.

The EEG data from the 64 channels of each participant trial were aligned to either of the two alignment points: the cue timing or GVS start. The EEG data were collected in different time periods relative to the alignment point. After removal of the baseline, the time course of the absolute EEG signal from each of the EEG channels (collected at 512 Hz) was used for the decoding. For example, in each trial for decoding data in a 96-ms time period relative to GVS (plotted in black at $x = 96$ ms in Fig. 2), the decoder was provided with $64 \text{ channels} \times 48 \text{ EEG data points after start of GVS} (\sim 96 \text{ ms at } 512 \text{ Hz}) = 3072$ features. We observed that taking the absolute of the signal was critical for good decoding performance, whereas inclusion of frequency band powers and temporal smoothing of the EEG signals were not.

For each participant, 80% of the trials from each class were selected randomly as training data for the decoder. Testing was performed on the remaining 20% trial data (from each class) to get a performance value. This 80–20 training–test check was repeated 20 times to get the performance statistics with each participant. The average and standard deviation across the 20 decodings for each participant are shown in Table 1.

Data plots

Decoding was performed for every time period for every participant. The across-participant decoder performance was plotted as box plots in Fig. 2. The red box plot at any abscissa point shows the decoding

performance (with the cue direction as label) with EEG features from the time period between the cue and the abscissa point. The black and pink box plots at any abscissa point are decoding performances, with MATCH/MISMATCH and cue direction labels, respectively, but using data from time periods between the GVS start time (which is labeled as zero in Fig. 2) and the abscissa point. The inset histograms show the participant ensemble decoding performance in the 20×7 participants = 140 tests, with data from each time period. The data from each participant are shown in a different color.

Figure 3 plots the decoding performance with the 96-ms time period. It replots cue (red and pink) and MATCH/MISMATCH (black) decoding data from Fig. 2. It also plots the decoding performance from our control experiment (green data), which was exactly the same as the main experiment, except for the fact that the GVS was applied across the forehead and back of neck of the participants.

Figure 4A plots the weights to each EEG channel assigned by the MATCHED/MISMATCHED decoder in the 96-ms time period, averaged across the 48 time points (corresponding to ~ 96 ms) and across all the participants. Note that the sparse logistic algorithm ensured that the weights of irrelevant (to the decoding) features are set to zero. The channels that contained nonzero weights in six or more of the seven participants were marked with a blue disk in Fig. 4A. We recorded the median of the time at which each channel was selected for each participant. Figure 4B plots the range (minimum and maximum) of the medians across the participants.

Finally, Fig. 5 plots the across-participant median, the 75% percentile, and the range (whiskers) of the participant reports of any vestibular sensation, muscle twitch, and tactile stimulation in our experiment to validate the GVS was subliminal. Participants scored any sensation they experienced, between 1 and 6 (0 representing no report).

REFERENCES AND NOTES

1. J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, T. M. Vaughan, Brain-computer interfaces for communication and control. *Clin. Neurophysiol.* **113**, 767–791 (2002).
2. A. Shakeel, M. S. Navid, M. N. Anwar, S. Mazhar, M. Jochumsen, K. Niazi, A review of techniques for detection of movement intention using movement-related cortical potentials. *Comput. Math. Methods Med.* **2015**, 346217 (2015).
3. G. R. Müller-Putz, A. Schwarz, J. Pereira, P. Ofler, From classic motor imagery to complex movement intention decoding: The noninvasive Graz-BCI approach. *Prog. Brain Res.* **228**, 39–70 (2016).
4. G. Townsend, V. Platsko, Pushing the P300-based brain-computer interface beyond 100 bpm: Extending performance guided constraints into the temporal domain. *J. Neural. Eng.* **13**, 026024 (2016).
5. M. Middendorf, G. McMillan, G. Calhoun, K. S. Jones, Brain-computer interfaces based on steady-state visual evoked response. *IEEE Trans. Rehabil. Eng.* **8**, 211–214 (2000).
6. L. O'Hare, Steady-state VEP responses to uncomfortable stimuli. *Eur. J. Neurosci.* **45**, 410–422 (2016).
7. K. Dremstrup, Y. Gu, O. F. D. Nascimento, D. Farina, Movement-related cortical potentials and their application in brain-computer interfacing, in *Introduction to Neural Engineering for Motor Rehabilitation*, D. Farina, W. Jensen, M. Akay, (Wiley-IEEE Press, 2013), pp. 253–266.
8. N. Birbaumer, A. Kübler, N. Ghanayim, T. Hinterberger, J. Perelmouter, J. Kaiser, I. Iversen, B. Kotchoubey, N. Neumann, H. Flor, The thought translation device (TTD) for completely paralyzed patients. *IEEE Trans. Rehabil. Eng.* **8**, 190–193 (2000).
9. N. Birbaumer, T. Elbert, A. G. M. Canavan, B. Roch, Slow potentials of the cerebral cortex and behavior. *Physiol. Rev.* **70**, 1–41 (1990).
10. H. Shibasaki, M. Hallett, What is the Bereitschaftspotential? *Clin. Neurophysiol.* **117**, 2341–2356 (2006).
11. L. A. Miner, D. J. McFarland, J. R. Wolpaw, Answering questions with an electroencephalogram-based brain-computer interface (BCI). *Arch. Phys. Med. Rehabil.* **79**, 1029–1033 (1998).

Table 1. % MATCH/MISMATCH decoding accuracy in the 96-ms time period.

| Participant | Median | Mean | SD |
|-------------|--------|-------|------|
| 1 | 89.58 | 89.27 | 5.86 |
| 2 | 72.50 | 72.00 | 5.71 |
| 3 | 79.17 | 77.11 | 5.74 |
| 4 | 87.50 | 87.19 | 4.18 |
| 5 | 97.92 | 97.71 | 2.02 |
| 6 | 77.08 | 78.33 | 6.02 |
| 7 | 88.54 | 87.50 | 5.97 |

12. J. R. Wolpaw, H. Ramoser, D. J. McFarland, G. Pfurtscheller, EEG-based communication: Improved accuracy by response verification. *IEEE Trans. Rehabil. Eng.* **6**, 326–333 (1998).
13. D. M. Wolpert, J. R. Flanagan, Motor prediction. *Curr. Biol.* **11**, R729–R732 (2001).
14. G. Ganesh, R. Osu, E. Naito, Feeling the force: Returning haptic signals influence effort inference during motor coordination. *Sci. Rep.* **3**, 2648 (2013).
15. E. Todorov, M. I. Jordan, Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.* **5**, 1226–1235 (2002).
16. S. J. Blakemore, C. D. Frith, D. M. Wolpert, Spatio-temporal prediction modulates the perception of self-produced stimuli. *J. Cogn. Neurosci.* **11**, 551–559 (1999).
17. Y.-w. Tseng, J. Diedrichsen, J. W. Krakauer, R. Shadmehr, A. J. Bastian, Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *J. Neurophysiol.* **98**, 54–62 (2007).
18. A. Takagi, G. Ganesh, T. Yoshioka, M. Kawato, E. Burdet, Physically interacting individuals estimate the partner's goal to enhance their movements. *Nat. Hum. Behav.* **1**, 0054 (2017).
19. T. Ikegami, G. Ganesh, Watching novice action degrades expert's performance: Causation between action production and outcome prediction of observed actions by humans. *Sci. Rep.* **4**, 6989 (2014).
20. T. Ikegami, G. Ganesh, Shared mechanisms in the estimation of self-generated actions and prediction of other's actions by humans. *eNeuro* (2017).
21. D. M. Wolpert, K. Doya, M. Kawato, A unifying computational framework for motor control and social interaction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **358**, 593–602 (2003).
22. L. Aymerich-Franch, D. Petit, A. Kheddar, G. Ganesh, Forward modelling the rubber hand: Illusion of ownership reorganizes the motor-sensory predictions by the brain. *R. Soc. Open Sci.* **3**, 160407 (2016).
23. X. Tian, D. Poeppel, Mental imagery of speech and movement implicates the dynamics of internal forward models. *Front. Psychol.* **1**, 166 (2010).
24. R. Gentili, C. E. Han, N. Schweighofer, C. Papaxanthis, Motor learning without doing: Trial-by-trial improvement in motor performance during mental training. *J. Neurophysiol.* **104**, 774–783 (2010).
25. T. Maeda, H. Ando, M. Sugimoto, Virtual acceleration with galvanic vestibular stimulation in a virtual reality environment, in *Proceedings of IEEE Virtual Reality, (VR 2005)*, Bonn, Germany, 12 to 16 March 2005).
26. R. C. Fitzpatrick, J. Marsden, S. R. Lord, B. L. Day, Galvanic vestibular stimulation evokes sensations of body rotation. *Neuroreport* **13**, 2379–2383 (2002).
27. R. C. Fitzpatrick, B. L. Day, Probing the human vestibular system with galvanic stimulation. *J. Appl. Physiol.* **96**, 2301–2316 (2004).
28. O. Yamashita, M. A. Sato, T. Yoshioka, F. Tong, Y. Kamitani, Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns. *Neuroimage* **42**, 1414–1429 (2008).
29. G. Pfurtscheller, F. H. Lopes da Silva, Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clin. Neurophysiol.* **110**, 1842–1857 (1999).
30. G. Pfurtscheller, C. Neuper, D. Flotzinger, M. Pregenzer, EEG-based discrimination between imagination of right and left hand movement. *Electroencephalogr. Clin. Neurophysiol.* **103**, 642–651 (1997).
31. R. Näätänen, A. W. K. Gaillard, S. Mäntysalo, Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol. (Amst)* **42**, 313–329 (1978).
32. G. Stefanics, P. Astikainen, I. Czigler, Visual mismatch negativity (vMMN): A prediction error signal in the visual modality. *Front. Hum. Neurosci.* **8**, 1074 (2014).
33. R. Näätänen, P. Paavilainen, T. Rinne, K. Alho, The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clin. Neurophysiol.* **118**, 2544–2590 (2007).
34. O. Korzyukov, K. Alho, A. Kujala, V. Gumenyuk, R. J. Ilmoniemi, J. Virtanen, R. Näätänen, Electromagnetic responses of the human auditory cortex generated by sensory-memory based processing of tone-frequency changes. *Neurosci. Lett.* **276**, 169–172 (1999).
35. A. Weinberg, R. Dieterich, A. Riesel, Error-related brain activity in the age of RDoC: A review of the literature. *Int. J. Psychophysiol.* **98**, 276–299 (2015).
36. C. B. Holroyd, M. G. H. Coles, The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* **109**, 679–709 (2002).
37. Y. Wang, Y. T. Wang, T. P. Jung, Translation of EEG spatial filters from resting to motor imagery using independent component analysis. *PLOS ONE* **7**, e37665 (2012).
38. G. H. Mulliken, S. Musallam, R. A. Andersen, Forward estimation of movement state in posterior parietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 8170–8177 (2008).
39. M. Desmurget, C. M. Epstein, R. S. Turner, A. C. Prablanc, G. E. Alexander, S. T. Grafton, Role of the posterior parietal cortex in updating reaching movements to a visual target. *Nat. Neurosci.* **2**, 563–567 (1999).
40. A. Reichenbach, J. P. Bresciani, H. H. Bühlhoff, A. Thielscher, Reaching with the sixth sense: Vestibular contributions to voluntary motor control in the human right parietal cortex. *Neuroimage* **124**, 869–875 (2016).
41. I. SanMiguel, K. Saupé, E. Schröger, I know what is missing here: Electrophysiological prediction error signals elicited by omissions of predicted “what” but not “when”. *Front. Hum. Neurosci.* **7**, 407 (2013).
42. S. Hirose, I. Nambu, E. Naito, An empirical solution for over-pruning with a novel ensemble-learning method for fMRI decoding. *J. Neurosci. Methods* **239**, 238–245 (2015).
43. E. Naito, T. Nakashima, T. Kito, Y. Aramaki, T. Okada, N. Sadato, Human limb-specific and non limb-specific brain representations during kinesthetic illusory movements of the upper and lower extremities. *Eur. J. Neurosci.* **25**, 3476–3487 (2007).
44. B. M. London, L. R. Jordan, C. R. Jackson, L. E. Miller, Electrical stimulation of the proprioceptive cortex (area 3a) used to instruct a behaving monkey. *IEEE Trans. Neural Syst. Rehabil. Eng.* **16**, 32–36 (2007).
45. J. Höhne, E. Holz, P. Staiger-Sälzer, K.-R. Müller, A. Kübler, M. Tangermann, Motor imagery for severely motor-impaired patients: Evidence for brain-computer interfacing as superior control solution. *PLOS ONE* **9**, e104854 (2014).
46. K. Oppenlander, K. S. Utz, S. Reinhart, I. Keller, G. Kerkhoff, A. K. Schaad, Subliminal galvanic-vestibular stimulation recalibrates the distorted visual and tactile subjective vertical in right-sided stroke. *Neurophysiologia* **74**, 178–183 (2015).

Acknowledgments

Funding: G.G. was partially supported by the Japan Science and Technology Agency (JST) Exploratory Research for Advanced Technology (ERATO), Japan (grant number JPMJER1701). Y.K. and Y.A. were supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI grants 17H05903 and 26112004. Part of this work was supported by JSPS KAKENHI (grant number 15K01849) and JST PRESTO (Precursory Research for Embryonic Science and Technology) (grant number JPMJPR17JA). **Author contributions:** G.G. developed the concept. G.G., Y.K., H.A., and E.Y. developed the experimental design. G.G., K.N., S.S., and A.M.T. collected the data. G.G., N.K., N.Y., and Y.K. analyzed the data. G.G., N.Y., and Y.K. wrote the paper. **Competing interests:** G.G., H.A., E.Y., and Y.K. are inventors on a pending patent application related to this work (application no. PCT/FR2016/51813, filed on 13 July 2016). The authors declare that they have no other competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper. Additional data related to this paper may be requested from the authors.

Submitted 22 September 2017

Accepted 26 March 2018

Published 9 May 2018

10.1126/sciadv.aqa0183

Citation: G. Ganesh, K. Nakamura, S. Saetia, A. M. Tobar, E. Yoshida, H. Ando, N. Yoshimura, Y. Koike, Utilizing sensory prediction errors for movement intention decoding: A new methodology. *Sci. Adv.* **4**, eaaq0183 (2018).