Error Correction Regression Framework for Enhancing the Decoding Accuracies of Ear-EEG Brain–Computer Interfaces

No-Sang Kwak and Seong-Whan Lee, *Fellow, IEEE*

***Abstract*—Ear-electroencephalography (EEG) is a promising tool for practical brain–computer interface (BCI) applications because it is more unobtrusive, comfortable, and mobile than a typical scalp-EEG system. However, an ear-EEG has a natural constraint of electrode location (e.g., limited in or around the ear) for acquiring informative brain signals sufficiently. Achieving reliable performance of ear-EEG in specific BCI paradigms that do not utilize brain signals on the temporal lobe around the ear is difficult. For example, steady-state visual evoked potentials (SSVEPs), which are mainly generated in the occipital area, have a significantly attenuated and distorted amplitude in ear-EEG. Therefore, preserving the high level of decoding accuracy is challenging and essential for SSVEP BCI based on ear-EEG. In this paper, we first investigate linear and nonlinear regression methods to increase the decoding accuracy of ear-EEG regarding SSVEP paradigm by utilizing the estimated target EEG signals on the occipital area. Then, we investigate an ensemble method to consider the prediction variability of the regression methods. Finally, we propose an error correction regression (ECR) framework to reduce the prediction errors by adding an additional nonlinear regression process (i.e., kernel ridge regression). We evaluate the ECR framework in terms of single session, session-to-session transfer, and subject-transfer decoding. We also validate the online decoding ability of the proposed framework with a short-time window size. The average accuracies are observed to be 91.11±9.14%, 90.52±8.67%, 86.96±12.13%, and 78.79±12.59%. This paper demonstrates that SSVEP BCI based on ear-EEG can achieve reliable performance with the proposed ECR framework.**

***Index Terms*—Brain–computer interface (BCI), ear-electroencephalography (EEG), nonlinear regression, steady-state visual evoked potential (SSVEP).**

Manuscript received November 20, 2018; revised March 21, 2019 and May 26, 2019; accepted June 14, 2019. Date of publication July 10, 2019; date of current version July 10, 2020. This work was supported in part by the Samsung Research Funding Center of Samsung Electronics under Project SRFC-TC1603-02, and in part by the Institute for Information & Communications Technology Planning & Evaluation grant funded by the Korea Government (Development of BCI-Based Brain and Cognitive Computing Technology for Recognizing User's Intentions Using Deep Learning) under Grant 2017-0-00451. This paper was recommended by Associate Editor C.-T. Lin. *(Corresponding author: Seong-Whan Lee.)*

N.-S. Kwak is with the Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, South Korea (e-mail: nskwak@korea.ac.kr).

S.-W. Lee is with the Department of Artificial Intelligence, Korea University, Seoul 02841, South Korea, and also with the Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, South Korea (e-mail: sw.lee@korea.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2019.2924237

I. Introduction

BRAIN–COMPUTER interfaces (BCIs) allow for interpreting and performing (a limited number of) user intentions using only brain signals [1], [2]. BCI users can harness brain states for controlling external devices (e.g., robot [3], wheelchairs [4], [5], and speller system [6]). Among the various brain imaging techniques (i.e., fMRI [7], fNIRS [8], and MEG [9]), electroencephalography (EEG) is a well-established technique that provides valuable insights into brain activities [10]–[13]. To efficiently generate user intentions, the existing EEG-based BCI techniques generally use BCI paradigms, such as motor imagery (MI) [14]–[16], P300 [17]–[20], and steady-state visual- or auditory-evoked potential (SSVEP [21]–[23] and SSAEP [24], [25], respectively).

Recent EEG-based BCI studies have demonstrated their applicability in real life from the perspective of machine learning as well as a portable hardware system and non-hair-bearing EEG (e.g., [26]–[34]). For example, ear-EEG [33], [34], which is a noninvasive electrical brain activity-measuring method using electrodes placed in or around the ear, has been increasingly developed due to its benefits over the conventional scalp-EEG system (i.e., convenience, unobtrusiveness, and mobility). The applications of ear-EEG include a hearing aid [35], [36]; sleep monitoring [37], [38]; biometric identification [39]; epilepsy detection system [40]; and fatigue estimation [41]. Furthermore, its availability with auditory BCI paradigms, including auditory event-related potential (ERP) [33], [42]–[44] and auditory steady-state response [34], [42], [43], [45]–[47], has been investigated. Auditory BCI with ear-EEG has shown reliable performance because it can directly achieve a relatively clear EEG signal near the temporal cortex.

Furthermore, an SSVEP paradigm combined with the ear-EEG has shown great potential for practical BCI systems [34], [41]–[43], [45]–[50]. SSVEPs are periodic responses elicited by the repetitive fast presentation of visual stimuli; they typically operate at frequencies between approximately 1 and 100 Hz, and they can be distinguished by their characteristic composition of harmonic frequencies [51]. In addition, BCI users can easily activate the SSVEPs according to their intentions by concentrating their gaze on the specific visual stimulus. However, the SSVEP BCIs based on ear-EEG have difficulty in accomplishing scalp-EEG-like performance because SSVEP strongly occurs in the occipital cortex. Note

that a proper channel location is an essential factor for a high decoding performance. Hence, the natural constraint of electrode location is an inevitable limitation of ear-EEG.

Previous studies have inspected the deterioration of the SSVEP quality in various ear-EEG electrodes. For example, Kidmose *et al.* [42] proposed an earplug-type ear-EEG electrode. Three classes (10, 15, and 20 Hz) of SSVEP performance, such as SNR, were measured in comparison to scalp-EEG. On average, SSVEP qualities of ear-EEG at the first harmonic frequencies were decreased from 30 to 10 dB. Looney *et al.* [49] found the SSVEP performance to decrease by approximately 50% (i.e., capacity ratios for scalp- and ear-EEG based on the estimated SNR and independent of the stimulus presentation) using ear-EEG with two LED visual stimuli (i.e., 15 and 20 Hz). Looney *et al.* [34] and Goverdovsky *et al.* [45] developed an earpiece-type of in-ear EEG sensor. The response obtained from the in-ear EEG was weaker than those obtained from the central regions. Wang *et al.* [50] compared SSVEP of the occipital area with that of the behind-ear area using a high-density EEG cap. As a result, the ear-EEG exhibited a performance reduction of approximately 50% compared to that of scalp-EEG. Furthermore, in other ear-EEG research, such as a carbon nanotube polydimethylsiloxane-based earphone-type EEG [43], high-density ear-EEG [46], and dry-contact electrode ear-EEG [47], a decreased SSVEP amplitude compared to that of scalp-EEG was reported.

Hence, this paper aims to enhance the deteriorated ear-EEG-based SSVEP decoding performance. Our main concept is to utilize the estimated EEG signals on the occipital area based on the linear and nonlinear regression models between ear-EEG and scalp-EEG signals. Mikkelsen *et al.* [52] proved that large portions of the scalp-EEG can be predicted from ear-EEG in the *alpha attenuation, auditory onset, and mismatch-negativity response paradigms*. They also formulated that the relation between ear-EEG and scalp-EEG was *linear* based on the *mutual information theory* [53]; note that the predictions based on the mutual information theory eventually led to the multiple linear regression (MLR) method used in our experiment.

However, usually, the prediction accuracy of the linear regression model is degraded by the noise of the independent variables in the training data. To minimize this effect, regularization techniques (e.g., l_1- and l_2-norm regularization) can be used by adding the shrinkage quantity. Furthermore, the linear method is a rough approximation of the real model because many complex and interesting phenomena in nature occur owing to the nonlinear phenomena. Furthermore, in research related to BCIs, regularization, and nonlinear techniques, such as least absolute shrinkage and selection operator (LASSO) [54], ridge regression (RR) [55], and kernel RR (KRR) [56], are employed to demonstrate a more proper interpretation and better results because EEG exhibits characteristics, such as noisiness, nonstationarity, multiscality, and high dimensionality [57].

For more accurate and reliable predictions, the nonlinearity property can be considered. Hence, to the best of our knowledge, this paper is the first attempt to enhance the ear-EEG performance in the *SSVEP paradigm* and utilize the *nonlinear relation* between ear-EEG and scalp-EEG. Furthermore, we are the first to propose an error correction regression (ECR) framework for SSVEP BCI based on ear-EEG, which can estimate and correct the errors by adding an additional nonlinear regression process. In addition, this is the first *online study* to demonstrate that nonlinear prediction can be used for real-time ear-EEG BCIs.

For this purpose, we investigate the signal estimation methods based on the regression methods (i.e., MLR, RR, and KRR); here, we utilize l_2-norm regularization because l_1-norm regularization generally causes the corresponding coefficient to tend to zero when a high correlation exists between independent variables (e.g., ear-EEG signals in very adjacent locations). We also investigate an ensemble regression (ER)-based method to consider the prediction variability of the above methods. Finally, we propose an ECR framework that has an additional nonlinear regression process (based on KRR) for estimating and correcting the errors of the previous regression process. While the concept of correcting the prediction errors with machine learning is not new [58], [59], we correct the errors by estimating the errors based on the nonlinear regression method by only using the ear-EEG system without the scalp-EEG system. We evaluate the proposed framework extensively in terms of a single session decoding, session-to-session transfer decoding, subject-transfer decoding, and online decoding.

The remainder of this paper is organized as follows. In Section II, the experimental setup, data acquisition, signal processing, and signal estimation methods and evaluation are described. The numerical results and discussion are presented in Section III and the conclusions and future work are outlined in Section IV.

II. Materials and Methods

A. Experimental Setup

Eleven healthy subjects with normal or corrected-to-normal vision and no history of neurological disease participated in this paper (age range: 25–32 years; 1 female). We simultaneously recorded the EEG signals using two devices. To acquire EEG signals in the occipital area, we used a BrainAmp (Brain Products GmbH) with eight channels (PO7, PO3, POz, PO4, PO8, O1, Oz, and O2); reference and ground electrodes were placed at Fpz and FCz, respectively. Here, we call the EEG signals scalp-EEG. To acquire ear-EEG signals, we used the Smarting System (mBrainTrain LLC) and cEEGrid electrodes with 18 channels; reference and ground electrodes were placed in the middle of the right ear cEEgrid (for a detailed description of the system and electrode, see [33]). Both devices applied a sampling frequency of 500 Hz, band-pass filter at 0.3–50 Hz, and a 60-Hz notch filter. We attached two cEEGrid electrodes around both ears and then put an EEG cap on the subject's head; Fig. 1 shows detailed channel information and the EEG device setup.

To induce SSVEP signals, we designed three flickering visual stimuli on the LCD monitor (Samsung, SyncMaster 2494HM). Fig. 2(a) shows a visual stimulation; the size of the stimuli was 8 cm $\times$ 8 cm. Each stimulus had frequencies

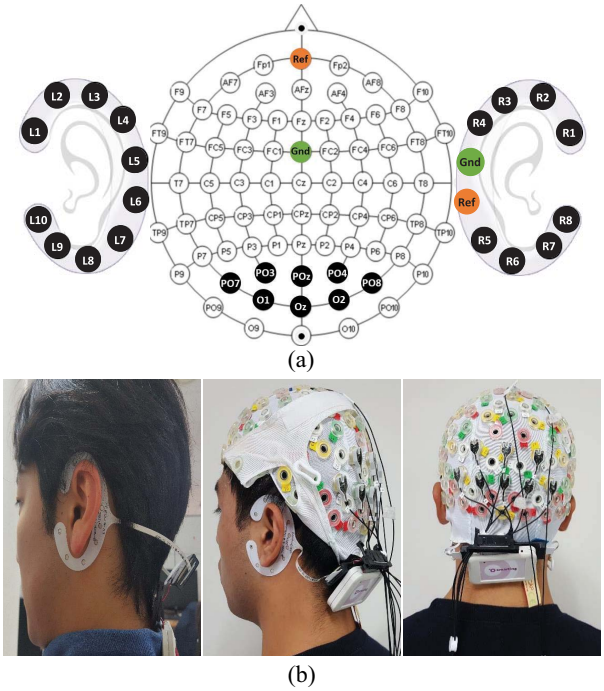

Fig. 1. (a) Channel locations of scalp-EEG and ear-EEG. (b) Appearance of wearing the EEG devices.

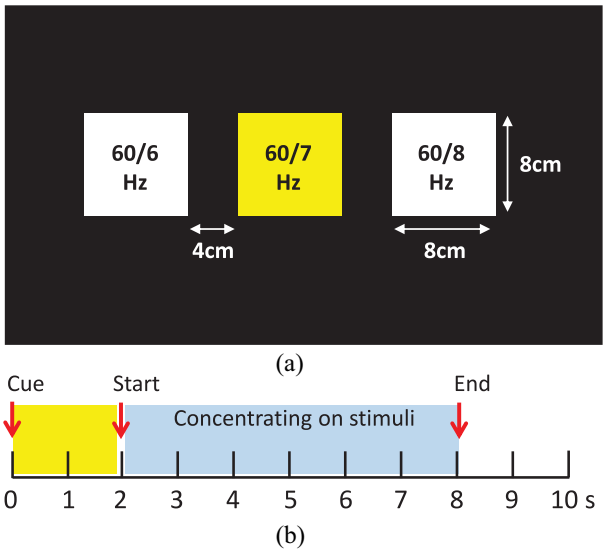

Fig. 2. (a) Size, shape, and flickering frequencies of the three visual stimuli. (b) Experimental procedure for sessions 1 and 2.

of 10, 8.57, and 7.5 Hz, which were calculated by dividing the monitor refresh rate by an integer (i.e., 60/6, 60/7, and 60/8). The visual stimuli were presented to the subjects using the Psychophysics Toolbox in MATLAB. During the data acquisition, subjects sat in a comfortable chair, and the distance between the subjects and the screen was maintained at approximately 60 cm.

B. Data Acquisition

We acquired three sessions of SSVEP data. In the first and second sessions (called session 1 and session 2, respectively), we used both the scalp-EEG and ear-EEG devices. In these sessions, we collected 150 trials in total (50 trials for each class). The subjects were instructed to attend, randomly, to the specific visual stimuli. Visual and auditory cues were given simultaneously. The auditory cue was provided to capture the attention and concentration of the subjects. The visual cue indicated the target class over 2 s with a yellow box on the target stimulus. The subjects were then asked to gaze at the corresponding stimulus for 6 s; during this time, all stimuli blinked simultaneously [Fig. 2(b)]. In the third session (session 3), we acquired only the ear-EEG signals without using scalp-EEG devices. We collected 60 trials in total (20 trials for each class). The instructions given to the subjects were similar to those in sessions 1 and 2, except that we requested them to concentrate on the stimulus for 3 s. Note that the three sessions were acquired on different days. The dataset (MATLAB file format) including the three sessions can be downloaded at http://deepbci.korea.ac.kr.

C. Signal Processing

From the acquired EEG signals, we applied the fourth Butterworth band-pass filter of [3–25] Hz to both scalp-EEG and ear-EEG signals (note that we did not use other artifact removal methods). To model the relation between the scalp-EEG (dependent variables) and the ear-EEG (independent variables), we modified the independent variables from the original EEG signals using the time-delay samples (τ). We constructed the temporally embedded ear-EEG ($\tilde{X}$) by concatenating the past time samples of each channel. The size of the time-delay samples was selected before training ECR framework; in this paper, we choose six representative time-delay samples ($\tau = 0$, 1, 4, 9, 49, and 99) to show how the decoding trend changes. The dimensions of the independent variables at a specific time were determined by [($\tau + 1$) samples $\times$ number of channels in ear-EEG]. For example, for $\tau = 99$, the dimension of the independent variables at a specific time was 1800 (100 samples $\times$ 18 channels). Then, we analyzed the regression models between two EEG device signals and estimated the scalp-EEG signals using only the ear-EEG signals (detailed description of the estimation methods is given below in signal estimation methods). Finally, the predicted scalp-EEG signals were used for classification; here, we employed canonical correlation analysis (CCA)-based SSVEP classification method [60]–[62] to investigate the effect of the proposed method (note that CCA is a simple SSVEP detection method that does not require any training data for classification). CCA is a multivariable statistical method that finds a pair of linear combinations such that the correlation between two canonical variables H and Z is maximized (note that we use H and Z to prevent confusion with variables X and Y in the regression method). In SSVEP BCI research, CCA has been used to decompose the EEG signals to extract stimulation frequency-related information. CCA finds the weight vectors, W_h and W_z, which maximize the correlation between the canonical variants $h = H'W_h$ and $z = Z'W_z$, by solving

$$\max_{W_h, W_z} \rho(h, z) = \frac{E[h'z]}{\sqrt{E[h'h]E[z'z]}} = \frac{E[W_h'HZ'W_z]}{\sqrt{E[W_h'HH'W_h]E[W_z'ZZ'W_z]}}. \quad (1)$$

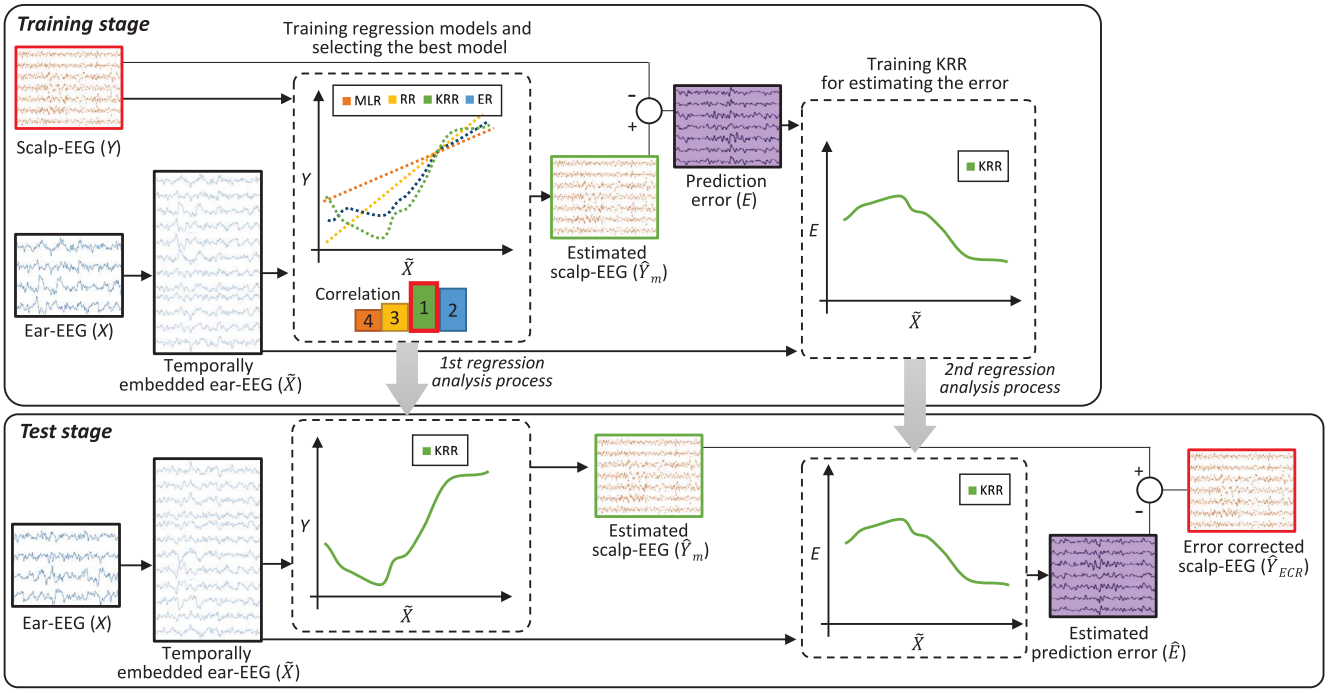

Fig. 3. Overview of the proposed ECR framework.

The maximum of ρ with respect to W_h and W_z is the maximum canonical correlation. As one dataset, we use EEG signals $H(t)$; here, $H(t)$ comprises the estimated EEG signals only or both the estimated EEG signals and ear-EEG signals. As $Z_{f_i}(t)$, we use the three reference frequencies (i.e., $f_1 = 10$, $f_2 = 8.57$, and $f_3 = 7.5$ Hz) of visual stimuli

$$Z_{f_i}(t) = (\sin(2\pi f_i t), \cos(2\pi f_i t), \sin(2\pi (2f_i)t), \cos(2\pi (2f_i)t))' \tag{2}$$

where t is the number of sampling points divided by the sampling rate. The maximized canonical correlation ρ is used for detecting the respective frequency of the stimulus that the subject attended. The classification result C is recognized by

$$C = \max_{f_i} \rho_{f_i}. \tag{3}$$

D. Signal Estimation Methods

Here, we describe five signal estimation methods. First, we explain three signal estimation methods based on the existing regression methods (i.e., MLR [63], RR [64], and KRR [65]). Subsequently, we present a signal estimation method based on the ensemble method ER of a linear combination of the three regression methods. Finally, we describe an ECR-based signal estimation method that can minimize the prediction error by using an additional regression process. The ECR-based signal estimation method is called ECR framework. Fig. 3 shows an overview of the proposed ECR framework.

1) MLR-Based: MLR is the most common form of linear regression analysis [63]. As a predictive analysis, MLR is used to explain the relation between continuous dependent variables and two (or more) independent variables. The predictions based on MLR, $\hat{Y}_{\mathrm{MLR}}$, can be written as

$$\hat{Y}_{\mathrm{MLR}} = B'_{\mathrm{MLR}} \tilde{X}. \tag{4}$$

The least-square error can be used to optimize weight B_{MLR} as

$$B_{\mathrm{MLR}} = \left(\tilde{X}\tilde{X}'\right)^{-1} \tilde{X} Y'. \tag{5}$$

In this paper, the dimension of independent variables $\tilde{X}$ (i.e., temporally embedded ear-EEG) is $(\tau+1)N_{\mathrm{ch_{ear}}} \times S$ and that of dependent variables Y is $N_{\mathrm{ch_{scalp}}} \times S$, where $N_{\mathrm{ch_{ear}}}$ is the number of channels in ear-EEG, $N_{\mathrm{ch_{scalp}}}$ is the number of channels in scalp-EEG, τ is the time-delay samples, and S is the number of time samples.

2) RR-Based: RR (linear least squares with l2-norm regularization) is a technique for analyzing multiple regression data that reduces the standard errors by adding a degree of bias to the regression estimates [64]. The prediction based on the RR method, $\hat{Y}_{\mathrm{RR}}$, can be written as

$$\hat{Y}_{\mathrm{RR}} = B'_{\mathrm{RR}} \tilde{X}. \tag{6}$$

The RR coefficient B_{RR} is

$$B_{\mathrm{RR}} = \left(\tilde{X}\tilde{X}' + \lambda I\right)^{-1} \tilde{X} Y' \tag{7}$$

where $\lambda \geq 0$ is a ridge parameter that controls the strength of the penalty term. Note that it becomes linear regression estimated using (5) as λ decrease to 0. On the other hand, increasing λ will shrink coefficients to 0. We used the MATLAB function *ridge* that returns a vector of coefficient estimates for a multilinear RR of the responses in Y on the predictors in $\tilde{X}$. In the training phase, we used ridge parameters ranging from 10^{-5} to 10^{-2} by generating ten points of logarithmically spaced vectors.

3) KRR-Based: KRR (i.e., the mean of a Gaussian process) combines RR with the kernel trick to estimate a nonlinear function of the input data [66]. It thus learns a linear function in the space induced by the respective kernel and the data. For nonlinear kernels, this corresponds to a nonlinear function in the original space. We can compute a kernel function $k(\cdot, \cdot)$ cheaply without ever explicitly forming a kernel feature space. It computes the inner product of two data points (p_i, p_j) in a kernel feature space ϕ. In this paper, we use a Gaussian kernel function

$$k(p_i, p_j) = e^{-(p_i - p_j)^2/\sigma} \tag{8}$$

where σ is the width of the Gaussian kernel function. The kernel trick essentially uses the fact that we can express the optimal nonlinear function as a linear combination of data similarities in a kernel feature space. For example, the prediction of KRR-based method $\hat{Y}_{\mathrm{KRR}}$ given a new data point $\tilde{x}_{\mathrm{new}}$ can be expressed as

$$\hat{Y}_{\mathrm{KRR}} = \left(\sum_j k(\tilde{x}_{\mathrm{new}}, \tilde{x}_j)\alpha_j \right)' \tag{9}$$

where j is an index variable that runs over all training data $\tilde{x}_j$ and α_j are dual coefficients obtained by

$$\alpha = (K + \lambda I)^{-1} Y' \tag{10}$$

where I is the identity matrix, λ denotes a ridge (regularization) parameter, and K is the square kernel matrix computed on the training data points. We used the ridge parameters ranging from 10^{-5} to 10^{-2} by generating five points of logarithmically spaced vectors. We also used a Gaussian kernel function with the width parameters ranging from 1 to 10^4 by generating ten points of logarithmically spaced vectors.

4) ER-Based: Prediction abilities of the above methods differ according to different time-delay samples (τ). For a stable prediction, we composed an ER consisting of the above three estimation methods. From the calibration phase in each method, averaged correlation values ρ_m (m indicates MLR-, RR-, and KRR-based signal estimation methods) between the estimated EEG signals and actual scalp-EEG were calculated using training and validation data. ER is a linear summation of the output of the three regression methods and is denoted by

$$\hat{Y}_{\mathrm{ER}} = \left(\frac{\rho_{\mathrm{MLR}} \hat{Y}_{\mathrm{MLR}} + \rho_{\mathrm{RR}} \hat{Y}_{\mathrm{RR}} + \rho_{\mathrm{KRR}} \hat{Y}_{\mathrm{KRR}}}{\rho_{\mathrm{MLR}} + \rho_{\mathrm{RR}} + \rho_{\mathrm{KRR}}} \right). \tag{11}$$

5) ECR-Based: ECR uses two regression processes. In the first process, it estimates $\hat{Y}_m$, where, m indicates MLR-based, RR-based, KRR-based, and ER-based signal estimation methods. We selected m that has a maximum of the averaged correlation value between the estimated EEG signals and the actual scalp-EEG, which was calculated using the training data. Furthermore, errors E between the estimated EEG signals $\hat{Y}_m$ and the actual EEG signals Y were obtained by

$$E = \hat{Y}_m - Y. \tag{12}$$

In the second regression process, we trained an additional regression model based on KRR between the new temporally embedded ear-EEG $\tilde{X}$ signals and the errors E of the previous regression process. Hence, the estimated errors $\hat{E}$ can be expressed as

$$\hat{E} = \left(\sum_l k(\tilde{x}_{\mathrm{new}}, \tilde{x}_l)\alpha_l \right)' \tag{13}$$

where l is an index variable that runs over all training data x_l and α_l is a dual coefficient obtained by

$$\alpha = (K + \lambda I)^{-1} E' \tag{14}$$

where I is the identity matrix, λ denotes a ridge parameter, and K is the square kernel matrix computed on all training data points. Note that we used different training data for the two regression processes (see details in Section II-E). The prediction based on the ECR framework, $\hat{Y}_{\mathrm{ECR}}$, can be written as

$$\hat{Y}_{\mathrm{ECR}} = \hat{Y}_m - \hat{E}. \tag{15}$$

We also used the Gaussian kernel function following (8) with the width parameters ranging from 1 to 10^4 and the ridge parameters ranging from 10^{-5} to 10^{-2} by generating ten points of logarithmically spaced vectors (note that hereafter we call the above methods MLR, RR, KRR, ER, and ECR, respectively).

E. Evaluation

We evaluated the proposed framework in terms of a single session decoding, session-to-session transfer decoding, subject-transfer decoding, and online decoding. For evaluating the single session decoding performance, sessions 1 and 2 were used for single session performance. In addition, we used a three-fold cross-validation to evaluate the models. We equally divided the whole data into training, validation, and test data to determine the parameters. Then, we performed a second (inner) cross-validation to determine the out-of-sample performance for a certain parameter configuration using the training and validation data; here, we used ridge parameters in the range 10^{-5}–10^{-2} and Gaussian kernel width parameters in the range 1–10^4 by generating ten points of logarithmically spaced vectors, respectively. This inner cross-validation was repeated for all parameter configurations, and the best configuration was used to train the algorithm using the training data. Finally, the trained model was tested using the test data. For training the model, we randomly selected 5% of the total number of samples from the training and validation data (i.e., 15 000 samples = 100 trials × 6 s × 500 samples/s × 5%); for testing the model, all samples of the test data were used (i.e., 150 000 samples = 50 trials × 6 s × 500 samples/s). For session-to-session transfer decoding, we used trained regression models in the single session evaluations. We selected the models that have the highest correlation value in the three-fold validation. We examined the four cases, from session 1 to session 2 (session 1→2), from session 1 to session 3 (session 1→3), from session 2 to session 1 (session 2→1), and from session 2 to session 3 (session 2→3). For subject-transfer decoding, we transferred the regression models of the best performing subject (i.e., Sub. 6) to others

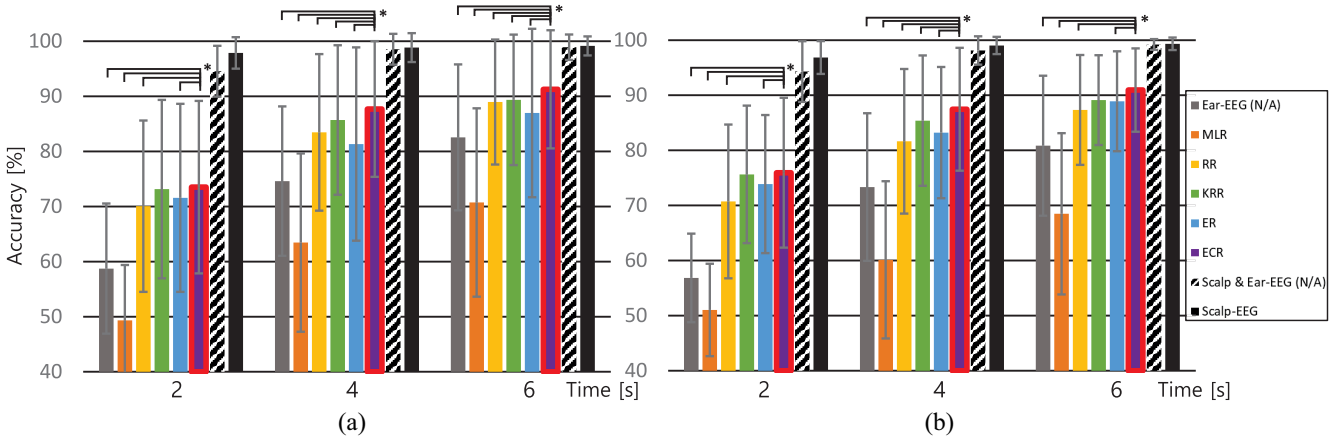

Fig. 4. Grand averaged accuracies of each method for 11 subjects according to the time window size (x-axis) using (a) session 1 and (b) session 2. Gray bar [i.e., ear-EEG (N/A)], black bar filled with lines [scalp-EEG and ear-EEG (N/A)], and black bar without lines (scalp-EEG) used actual measured signals. Colored bars show the accuracies from estimated signals without actual measured ear-EEG signals. Red boxes indicate the best accuracies among the performances using only ear-EEG. Error bars are standard deviations. One asterisk indicates 5% significance level between ECR (purple bar) and the corresponding method (Wilcoxon signed-rank test).

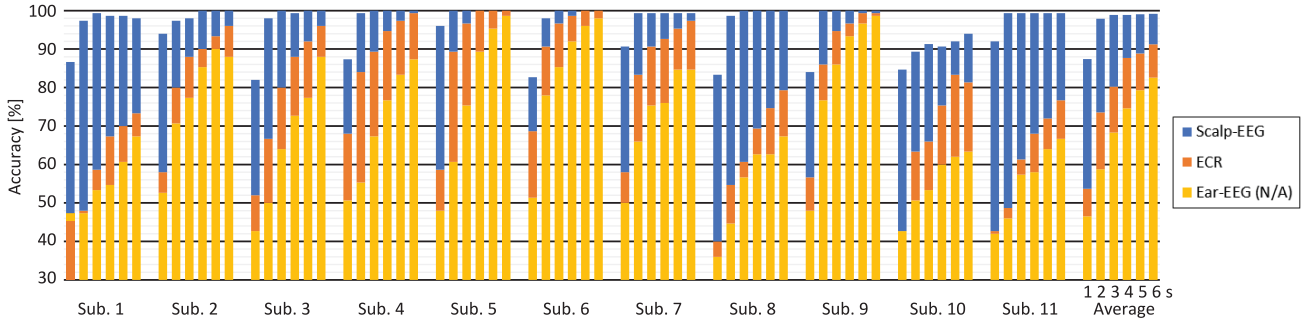

Fig. 5. Individual accuracies of ECR for single session 1. Six bars in a subject indicate six different window sizes from 1 s (left) to 6 s (right). Orange bars show increased accuracy compared to ear-EEG (yellow bar) where no signal estimation methods are applied. Blue bars indicate the accuracy of scalp-EEG.

on session 1 data (i.e., best→others). We also tested to transfer the model of worst performing subject (i.e., Sub. 1) to the others (i.e., worst→others). For online decoding, we tested the ECR method with only the ear-EEG system (without wearing the scalp-EEG) in a cue-based online manner using 2-s window size (30 trials in total). We also used trained regression models in the session-to-session transfer evaluation.

III. Results and Discussion

A. Single Session Decoding Performance

The single session decoding performance (when $\tau = 99$) is presented in Fig. 4, showing grand averages of each method for 11 subjects according to the three time window sizes (i.e., 2, 4, and 6 s). We observed more deteriorated performance of ear-EEG than scalp-EEG when the time window size was smaller when comparing the performance in measured scalp-EEG (black) and measured ear-EEG (gray), which did not apply any signal estimation method. The ECR showed the best accuracy in the different time windows (marked as red boxes in Fig. 4), also showing a significant difference with the other methods; statistical significance was set at 5% (Wilcoxon signed-rank test). Comparing the accuracies of ear-EEG when no method was applied, the averaged increments were 14.79±7.92, 13.09±7.02, and 8.71±6.43% (in session 1) at each time window size and 19.09±7.78, 14.12±4.75, and 10.10±5.83% (in session 2). The grand average in ECR of the two single sessions at 6 s was 91.11±9.14%.

Fig. 5 shows an individual performance increase using ECR according to the time window size in session 1. Although there are high performing subjects (i.e., Subs. 5, 6, and 9) in ear-EEG, performance decrease is inevitable compared to that of scalp-EEG and the other subjects showed significant deterioration.

The averaged correlation coefficient values between the measured scalp-EEG and estimated signals from ear-EEG based on ECR were recorded as 0.5447±0.0500 and 0.5638±0.0690 for sessions 1 and 2, respectively. Furthermore, correlation in ECR showed significant differences compared to those of the other methods (Wilcoxon signed-rank test, <0.05) (see details in Table I).

Fig. 6 shows an example of how the ECR method functions in the time domain; the result was randomly selected in session 1 for Sub. 5 at channel POz. In this example, KRR was selected for the first regression process in the training step; the correlation coefficient between the measured scalp-EEG and estimated signals by KRR was 0.3210. The errors between the two signals obtained by (12) were then measured. In addition, the second regression process estimated the

TABLE I

PERFORMANCE EVALUATION RESULTS ON SINGLE SESSION DECODING USING A THREE-FOLD CROSS-VALIDATION (WHEN $\tau = 99$). AVERAGED CORRELATION BETWEEN THE MEASURED SCALP-EEG AND THE ESTIMATED SIGNALS OF EACH METHOD WAS CALCULATED USING TEST DATA. AVERAGED ACCURACY WAS GIVEN AT THE 6 S TIME WINDOW. IN THE ACCURACY RESULTS, *Both* INDICATES CLASSIFICATION RESULTS USING BOTH ACTUAL MEASURED EAR-EEG SIGNALS AND ESTIMATED SIGNALS. AND ONLY INDICATES THAT OF USING ONLY ESTIMATED SIGNALS BASED ON THE CORRESPONDING METHOD. BOLD NUMBERS INDICATE THE BEST PERFORMING NUMBER. ASTERISK MARKS ON THE RESULTS INDICATE A SIGNIFICANT DIFFERENCE (WILCOXON SIGNED-RANK TEST, <0.05) BETWEEN ECR AND THE CORRESPONDING METHOD

Single Session	Session 1					Session 2				
Method	MLR	RR	KRR	ER	ECR	MLR	RR	KRR	ER	ECR
Correlation	0.0288 ±0.0157*	0.3690 ±0.0646*	0.5346 ±0.0539*	0.3501 ±0.0641*	**0.5447 ±0.0500**	0.0311 ±0.0139*	0.4087 ±0.0802*	0.5546 ±0.0742*	0.4196 ±0.1282*	**0.5638 ±0.0690**
Acc. (Both) [%]	82.97 ±13.29*	88.30 ±11.65*	89.31 ±11.03*	87.39 ±13.22*	**90.79 ±10.17**	81.52 ±12.24*	86.97 ±10.01*	89.45 ±8.89	87.21 ±9.81*	**90.42 ±8.61**
Acc. (Only) [%]	70.73 ±17.09*	88.97 ±11.34*	89.36 ±11.85*	86.97 ±15.29*	**91.26 ±10.72**	68.48 ±14.62*	87.33 ±9.97*	89.12 ±8.15	88.91 ±9.07*	**90.94 ±7.56**

TABLE II

PERFORMANCE EVALUATION RESULTS ON SESSION-TO-SESSION TRANSFER DECODING (WHEN $\tau = 99$). AVERAGED ACCURACY WAS GIVEN AT THE 6 S TIME WINDOW; SESSION 1 (TRAINING)→ SESSION 3 (TEST) AND SESSION 2 (TRAINING) → SESSION 3 (TEST) WERE EVALUATED WITH THE 3 S TIME WINDOW. AVERAGED CORRELATION BETWEEN THE MEASURED SCALP-EEG AND THE ESTIMATED SIGNALS OF EACH METHOD WAS CALCULATED USING TEST DATA (NOTE THAT WE ONLY MEASURED THE EAR-EEG SIGNALS IN SESSION 3). IN THE ACCURACY RESULTS, *Both* INDICATES CLASSIFICATION RESULTS USING BOTH ACTUAL MEASURED EAR-EEG SIGNALS AND ESTIMATED SIGNALS. AND ONLY INDICATES THAT OF USING ONLY ESTIMATED SIGNALS BASED ON THE CORRESPONDING METHOD. BOLD NUMBERS INDICATE THE BEST PERFORMING NUMBER. ASTERISK MARKS ON THE RESULTS INDICATE A SIGNIFICANT DIFFERENCE (WILCOXON SIGNED-RANK TEST, <0.05) BETWEEN ECR AND THE CORRESPONDING METHOD

Session Transfer	Session 1 (Training) → Session 2 (Test)					Session 2 (Training) → Session 1 (Test)				
Method	MLR	RR	KRR	ER	ECR	MLR	RR	KRR	ER	ECR
Correlation	0.0275 ±0.0437*	0.2702 ±0.1028*	0.5114 ±0.0672*	0.3234 ±0.1552*	**0.5255 ±0.0629**	0.0032 ±0.0270*	0.2433 ±0.1429*	0.4851 ±0.0900*	0.3058 ±0.1689*	**0.5059 ±0.0820**
Acc. (Both) [%]	83.94 ±11.95*	85.27 ±11.70*	89.76 ±8.34	86.30 ±11.26*	**89.82 ±8.37**	83.33 ±13.54*	85.45 ±13.01*	88.79 ±10.39*	86.67 ±11.67*	**90.12 ±9.80**
Acc. (Only) [%]	75.21 ±15.88*	82.73 ±14.71*	87.21 ±8.84*	84.12 ±14.16*	**90.30 ±7.93**	74.42 ±14.84*	81.76 ±15.08*	88.73 ±10.24*	84.67 ±12.79*	**90.73 ±9.42**
Session Transfer	Session 1 (Training) → Session 3 (Test)					Session 2 (Training) → Session 3 (Test)				
Method	MLR	RR	KRR	ER	ECR	MLR	RR	KRR	ER	ECR
Acc. (Both) [%]	71.97 ±15.66*	74.39 ±16.30*	81.06 ±12.64	76.51 ±15.17*	**81.36 ±13.31**	69.70 ±16.50*	76.67 ±15.88*	79.70 ±14.36	76.97 ±16.80*	**80.15 ±14.87**
Acc. (Only) [%]	61.67 ±15.29*	71.82 ±15.64*	80.76 ±12.41	73.18 ±16.48*	**82.27 ±10.04**	60.00 ±12.97*	76.21 ±16.67*	81.82 ±12.42*	78.48 ±14.52*	**83.33 ±11.85**

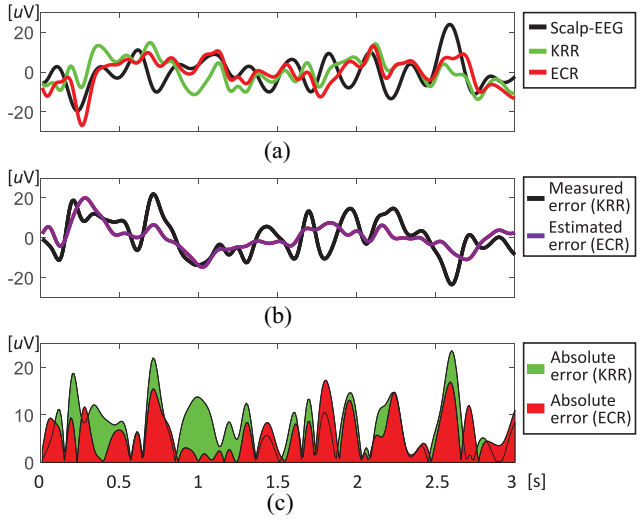

Fig. 6. (a) Example of the measured signals from the scalp-EEG (black) and the estimated signals from KRR (green) and ECR (red) in the time domain; the correlation of KRR (between the measured signals) is 0.3210 and ECR is 0.5660. (b) Measured error of KRR (black) at the same time with (a) and the estimated error by ECR (purple); the correlation is 0.5655. (c) Absolute error of estimated signals from KRR (green) and ECR (red) in (a); mean absolute error from KRR is 7.75 and that of ECR is 5.83. Note that this plot is randomly selected from the Sub. 5 at channel POz. The correlation value can differ.

error signals [Fig. 6(b)]; the correlation coefficient between the measured errors of KRR and the estimated errors by ECR was 0.5655. By subtracting the estimated errors from the estimated scalp-EEG signals, finally, the ECR output showed an enhanced correlation coefficient up to 0.5660. Fig. 6(c) shows a reduced estimation error by ECR compared to KRR.

B. Session-to-Session Transfer Decoding Performance

We trained the regression models in the specific session and tested them in completely new sessions within the same subject (i.e., intrasubject variability). The decoding performances (when $\tau = 99$) are given in Fig. 7 and Table II. In the case of session 1→2, 80.85% of ear-EEG accuracy, with no method applied, could be maximally enhanced up to 90.30% using ECR at the 6 s time window, and 2→1 increased 82.55% of ear-EEG to 90.73%. ECR showed the best accuracies in the different time windows. The grand average of two session-to-session transfer decoding at 6 s was 90.52±8.67%.

In the case of session 1→3 [Fig. 7(c)], 69.39±15.54% of ear-EEG was increased to 82.27±10.04%, and in the case of session 2→3 [Fig. 7(d)], increased up to 83.33±11.85% at the 3 s time windows (note that we acquired an SSVEP trial with the 3 s time window without scalp-EEG). ECR outperformed the others in 2 and 3 s time window; however, KRR had higher accuracy than the 1 s time window in session 1→3.

The outstanding ECR performance implies that the prediction errors could be transferred beyond the session within the subjects. In addition, the results with a small window size showed the applicability of a real-time system with the reliable performance of the proposed framework.

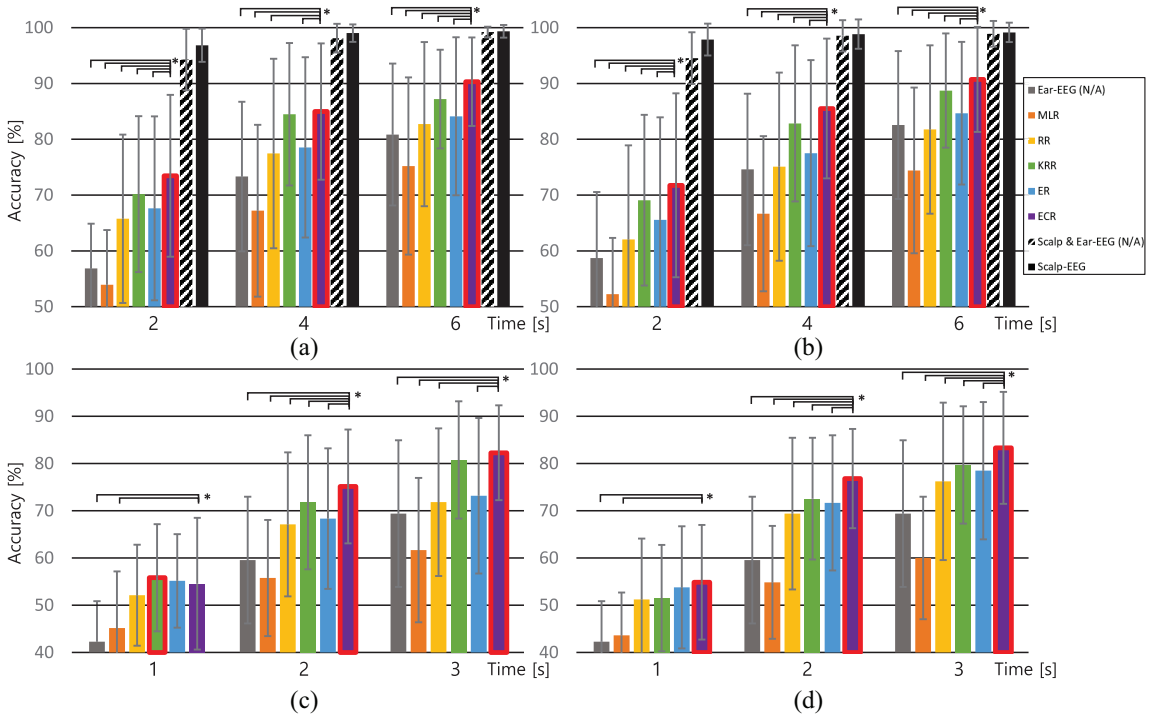

Fig. 7. Session-to-session transfer decoding performance of (a) session 1→2, (b) session 2→1, (c) session 1→3, and (d) session 2→3. One asterisk indicates 5% significance level between ear-EEG (gray bar) and the corresponding method (Wilcoxon signed-rank test).

C. Subject-Transfer Decoding Performance

We evaluated the subject-transfer decoding performances (i.e., intersubject variability) by transferring the best performing subject's regression models (best→others). We also found performance increases by ECR. However, as expected, the level of improvement was not sufficient for the above evaluations. The average accuracies of ECR were 68.44±14.22%, 81.59±14.31%, and 86.96±12.13% for the three time windows. Here, ECR sustained the significant differences (Fig. 8). We also tested transferring the worst performing subject's model (worst→others). In this case, ECR did not show the best accuracy (Table III); instead, KRR achieved the best performance. Furthermore, the tendency of statistical significance decreased compared to that of best→others. We assume that the worst subject's error correcting model in ECR was not suitable for other subjects to correct the prediction error. Note that building an optimal transfer model is an important issue and needs to be further investigated.

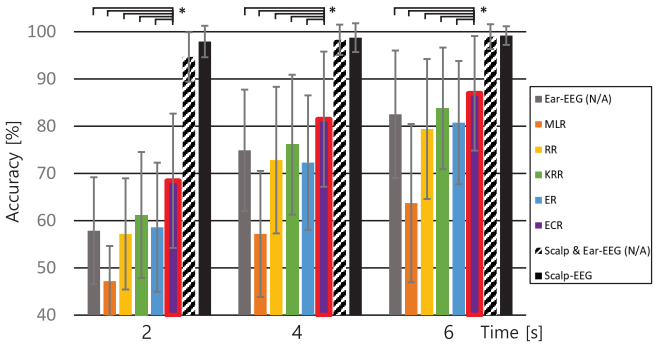

Fig. 8. Subject-transfer decoding performance on session 1: (a) best(Sub. 6)→others and (b) worst(Sub. 1)→others. Note that we excluded the results of Subs. 1 and 6 for an impartial comparison. The gray bar (ear-EEG), black bar filled with lines (scalp-EEG and ear-EEG), and black bar without lines (ear-EEG) used the actual measured signals. The other colored bars filled with lines show the accuracies acquired from *both* actual ear-EEG signals and estimated signals from the corresponding method. Colored bars without lines *only* show the accuracies from the estimated signals without actual ear-EEG signals. Red boxes indicate the best accuracies among the regression methods. Asterisk indicates 5% significance level between ear-EEG (gray bar) and the corresponding method (Wilcoxon signed-rank test).

D. Online Decoding Performance

The online decoding performance of ECR was evaluated using a 2-s window size. The results of accuracies and information transfer rate (ITR) are presented in Table IV. ITR was calculated as

$$\mathrm{ITR} = \left(\frac{60}{T}\right)\left(\log_2 N_c + A\log_2 A + (1-A)\log_2\left(\frac{1-A}{N_c-1}\right)\right) \tag{16}$$

where T is the time needed to classify user intention, N is the number of classes, and A is the accuracy. While T was 2.0015 s when ECR was not applied, ECR needed an average computational time T of 2.3079 s; we used a personal computer (Intel i5-3330 CPU@3-GHz RAM: 32 GB). ECR recorded an average accuracy of 78.79±12.59% and average ITR of 18.07±10.04 [bits/min]. Although ECR needed more computational time due to the two nonlinear processes, encouraging enhancement in accuracy and ITR demonstrates the feasibility of online use of the proposed method.

E. Spectral Analysis of ECR

Fig. 9 represents the averaged FFT plots within the same stimulus frequencies to compare the SSVEP power of the

TABLE III

PERFORMANCE EVALUATION RESULTS ON SUBJECT-TRANSFER DECODING (WHEN $\tau = 99$). AVERAGED CORRELATION BETWEEN THE MEASURED SCALP-EEG AND THE ESTIMATED SIGNALS OF EACH METHOD WAS CALCULATED USING TEST DATA. AVERAGED ACCURACY WAS GIVEN AT THE 6 S TIME WINDOW. IN THE ACCURACY RESULTS, *Both* INDICATES CLASSIFICATION RESULTS USING BOTH ACTUAL MEASURED EAR-EEG SIGNALS AND ESTIMATED SIGNALS. AND ONLY INDICATES THAT OF USING ONLY ESTIMATED SIGNALS BASED ON THE CORRESPONDING METHOD. BOLD NUMBERS INDICATE THE BEST PERFORMING NUMBER. RESULTS OF BEST PERFORMING SUBJECT (SUB. 6) AND WORST PERFORMING SUBJECT (SUB. 1) WERE EXCLUDED IN THE ANALYSIS. ASTERISK MARKS ON THE RESULTS INDICATE A SIGNIFICANT DIFFERENCE (WILCOXON SIGNED-RANK TEST, <0.05) BETWEEN ECR AND THE CORRESPONDING METHOD

Subject Transfer	Best Performing Sub. 6 (Training) → Others (Test)					Worst Performing Sub. 1 (Training) → Others (Test)				
Method	MLR	RR	KRR	ER	ECR	MLR	RR	KRR	ER	ECR
Correlation	0.0438 ±0.0143*	0.2532 ±0.1264*	0.4502 ±0.0545*	0.2146 ±0.1065*	**0.4562 ±0.0581**	0.0083 ±0.0199*	0.1909 ±0.1419*	0.4472 ±0.0792	0.2396 ±0.1155*	**0.4479 ±0.0818**
Acc. (Both) [%]	81.85 ±13.34*	85.19 ±12.93*	86.67 ±11.85*	86.22 ±12.10*	**88.44 ±11.43**	82.44 ±12.53*	84.52 ±12.30	**88.15 ±11.00**	85.85 ±12.19	86.67 ±13.58
Acc. (Only) [%]	63.70 ±16.76*	79.41 ±14.82*	83.78 ±12.87*	80.74 ±13.05*	**86.96 ±12.13**	60.15 ±15.23*	74.15 ±15.53*	**84.22 ±12.44**	77.56 ±77.56*	82.37 ±14.51

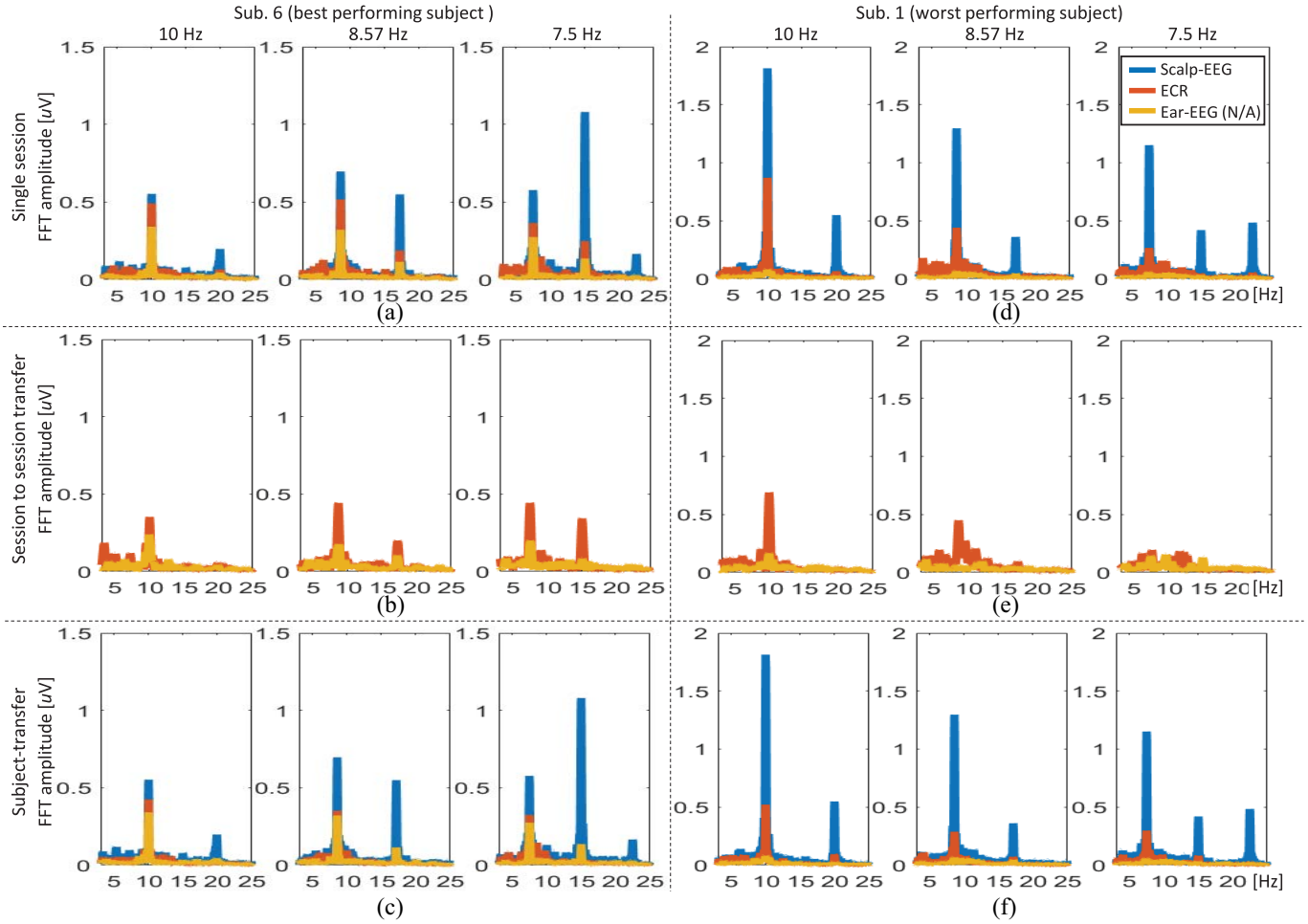

Fig. 9. Averaged FFT amplitude plots of all channels for Sub. 6 (a)–(c) and Sub. 1 (d)–(f). (a) and (d) Results from single session 1 and (b) and (e) session-to-session decoding (session 1→3). (c) and (f) Subject-transfer decoding. Blue, orange, and yellow lines indicate the scalp-EEG, estimated signals using ECR, and actual measured ear-EEG signals.

measured scalp-EEG signals (blue), ECR-based estimated EEG signals (orange), and measured ear-EEG signals (yellow) for Sub. 6 (best performing subject) and Sub. 1 (worst performing subject); we selected the two subjects based on the ECR accuracy at the 6 s time window. Most subjects with the scalp-EEG showed reliable accuracies, which were quite higher than the existing SSVEP research because we used only three target stimuli with relatively large size. Sub. 6 showed quite distinct SSVEP power in ear-EEG at the target frequency, whereas the peaks of Sub. 1 in ear-EEG were greatly smeared out. However, after the ECR was applied, the peaks were increased at the target frequencies in all evaluations; in the case of Sub. 6, the harmonic frequency component also increased. In the case of Sub. 1, there was also a power increase but not as high as that of Sub. 6. Thus, it is also difficult to find the effects at the harmonic frequencies.

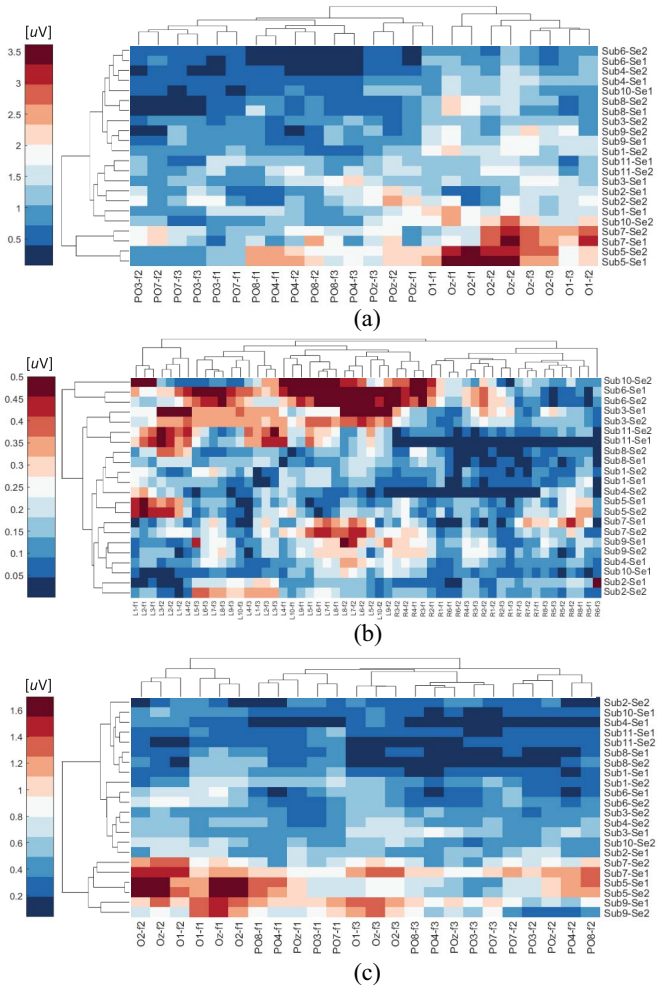

Fig. 10. Hierarchical cluster analysis using (a) scalp-EEG, (b) measured ear-EEG, and (c) ECR-based estimated EEG.

F. Hierarchical Cluster Analysis

The hierarchical cluster analysis, which seeks to build a hierarchy of clusters, can be used to extract and visualize the characteristics of EEG signals efficiently [67], [68]. We present the hierarchical cluster analysis results for three conditions (i.e., scalp-EEG [Fig. 10(a)], measured ear-EEG [Fig. 10(b)], and ECR-based estimated EEG [Fig. 10(c)]) to investigate the variability of SSVEP dynamics across multiple subjects and sessions. We constructed a heat map to illustrate the FFT amplitude for the stimuli frequency. The hierarchical cluster analysis was applied to both dimensions of the heat map to assess the intersubject and intrasubject variability (along rows) and the similarity among channels and stimuli frequency (along columns) with average-linkage dendrograms. We used Euclidian distance metric for estimating the similarity. Three subfigures showed similar result trends. Subjects with different sessions were grouped together, indicating a smaller intrasubject variability than intersubject variability. Further, adjacent channels with the same frequency were grouped together, indicating high similarity of the adjacent channels. However, it can be still inferred that the proposed ECR framework enhances the performance of ear-EEG; range of the FFT power in Fig. 10(c) was increased compared with that of Fig. 10(b).

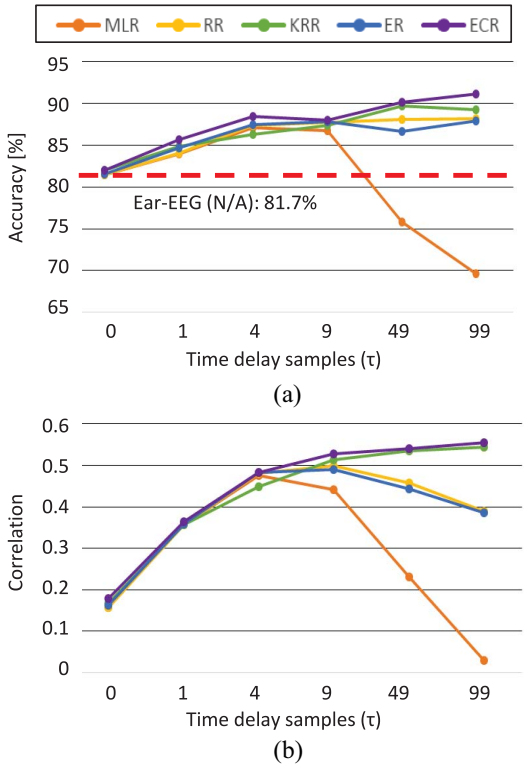

Fig. 11. Decoding trends of (a) accuracy and (b) correlation according to time-delay samples τ.

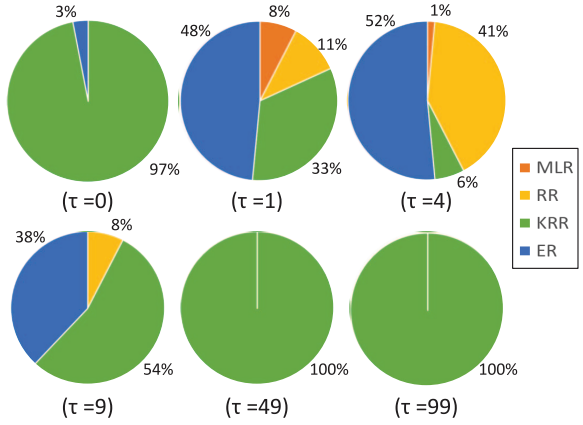

Fig. 12. Ratio of the selected regression method in ECR according to time-delay samples τ.

G. Decoding Trends

According to the number of time-delay samples (τ) for temporally embedded ear-EEG input features (i.e., different dimensions of input features), decoding results of each estimation method with a 6 s time window showed different trends in accuracy and correlation (Fig. 11). We tested different time-delay samples (e.g., $\tau = 0$, 1, 4, 9, 49, and 99) using data from sessions 1 and 2. In particular, performance degradation of MLR occurred gradually and then considerably when τ was larger than 9. On the other hand, KRR showed an increase in performance as τ grew. We conjecture that the nonlinear characteristics of the input features are prominent

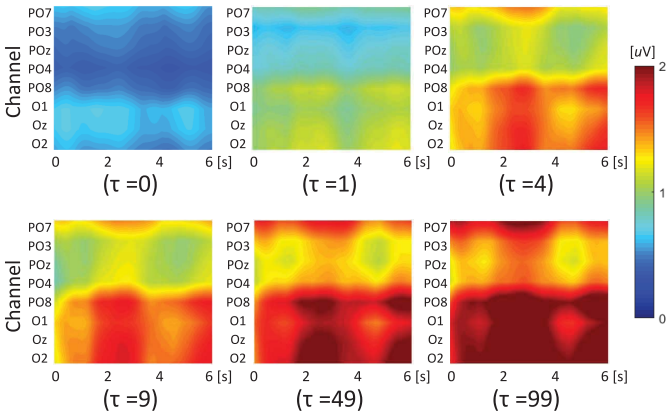

Fig. 13. Spatial and temporal maps according to time-delay samples τ for Sub. 8.

TABLE IV
ONLINE EXPERIMENTAL RESULTS: TARGET SEQUENCE WAS RANDOMLY GIVEN AMONG THE THREE CLASSES. N/A INDICATES CLASSIFICATION RESULT WHEN ECR WAS NOT APPLIED. PARENTHESIS INDICATES STANDARD DEVIATION

Sub.	Sequence		Acc.	ITR
1	Target	2213322213 3211132112 1333322113	-	-
	N/A	1211133312 1311121112 1111221112	43.33	0.93
	ECR	2112122211 1111122112 1331111113	63.33	7.02
2	Target	1222311121 1233231323 1322311332	-	-
	N/A	3222323122 2233333333 2222333332	60.00	6.42
	ECR	3222331323 3233233323 1222333332	70.00	10.49
3	Target	1313213331 1331213222 2221311322	-	-
	N/A	2112113111 1311111111 2121131113	50.00	1.64
	ECR	1333113331 1331113223 2223321123	73.33	12.52
4	Target	1333322132 2222212331 1323113111	-	-
	N/A	1333132132 2323232332 3122123112	63.33	8.10
	ECR	1333322132 2222232231 1223123311	83.33	19.97
5	Target	2121313132 1133221222 1223331313	-	-
	N/A	2121213332 1131223333 1312311133	56.67	4.93
	ECR	2121311112 1133221222 1223331313	93.33	30.28
6	Target	3313123212 2311122323 2332113112	-	-
	N/A	2321313211 1232132313 2332121222	46.67	1.64
	ECR	3313123212 1311122323 2332113112	96.67	34.86
7	Target	2121231213 2313321131 3112232332	-	-
	N/A	2121231111 2313211123 3232232331	70.00	12.10
	ECR	2121211213 2313321133 3122232332	90.00	26.41
8	Target	2331231311 1122332211 1323213322	-	-
	N/A	2331333111 1112221311 1332112123	53.33	3.64
	ECR	1331121111 3121322311 1322112323	60.00	5.56
9	Target	3321121133 3132322121 1233322121	-	-
	N/A	3322112123 3212332321 3112232121	56.67	4.93
	ECR	3321122123 3132333121 1232232311	70.00	10.49
10	Target	2212332112 3131123313 3132122321	-	-
	N/A	2212212321 3131311312 3123223221	53.33	3.64
	ECR	2222332112 2131123313 3122122321	90.00	26.41
11	Target	1123331113 3332231213 1232221212	-	-
	N/A	2112313112 1311233112 2221121221	40.00	0.42
	ECR	1122311112 3312231212 1232211211	76.67	14.77
Avg.	N/A		53.94 (8.80)	4.48 (3.43)
	ECR		78.79 (12.59)	18.07 (10.04)

as the value of τ increases. ER showed stable decoding trends until τ was 9. However, when τ was larger than 9, ER could not avoid the impact of the large performance degradation of MLR. ECR had the highest average accuracy and correlation regardless of τ; the correlation of ECR was slightly better than that of the others, which resulted in enhanced accuracy. These results show that the error correction process in ECR could enhance the decoding performance of ear-EEG-based SSVEP BCI. Fig. 12 provides the ratio of the selected regression methods (i.e., MLR, RR, KRR, and ER) in the first regression process in ECR. The total number of selected methods is 66 (3 cross-validations, 11 subjects, and 2 sessions). According to τ, the selected method differed. When τ was 0, 9, 49, and 99, KRR was the most frequently selected. On the other hand, ER was the most frequently selected when τ was 1 and 4. Fig. 13 shows the spatial and temporal maps of the ECR-based estimated EEG signals according to different time-delay samples. The SSVEP trials corresponding to f_1 (i.e., 10 Hz) class were transformed into the frequency domain using short-time Fourier transform with 1024 samples of window length, 8 samples of hop size, and 4096 samples of FFT points. Then, 10-Hz components were averaged at each time point. The averaged amplitude at 10 Hz for all channels shows an increase with τ. Further, three reconstructed channels in occipital area (i.e., O1, Oz, and O2) tend to have a larger value compared with reconstructed parietal–occipital channels (i.e., PO7, PO3, POz, PO4, and PO8).

IV. CONCLUSION

SSVEP combined with the scalp-EEG system has shown great potential for promoting the communication ability [69], assisting motor function [21], and controlling external devices [70], as well as in percepting the video quality [71]. From the perspective of BCI users, ear-EEG has a significant opportunity to realize the real-life use of BCI application because it is a more convenient, unobtrusive, and user-friendly device compared to the conventional scalp-EEG system. However, due to its constraint of electrode locations, it is difficult to achieve a reliable (scalp-EEG-like) performance with the various BCI paradigms. SSVEP in ear-EEG could be detected but with deteriorated signal quality through blurring of the field by the head volume conductor [72].

Hence, this paper has presented a signal estimation method for enhancing the performance of ear-EEG in SSVEP paradigms using linear and nonlinear regression models. Furthermore, we proposed a novel framework, ECR, that can make more sophisticated predictions by estimating and correcting the errors using the nonlinear regression method. Consequently, ECR increased the correlation coefficient values between the estimated and measured signals, which means an improvement in estimation accuracy. It thus significantly increased the performance of SSVEP classification abilities. We evaluated the proposed methods in terms of single session decoding, session-to-session transfer decoding, and subject-transfer decoding performance. In all cases, the ECR framework showed the best increases in SSVEP based on ear-EEG. In particular, the performance improvement in a single session was the highest. We also find that our proposed framework could deal with both session-to-session and subject-to-subject variability, achieving higher accuracy compared to using only ear-EEG signals; however, building an optimal transfer model needs to be further investigated based on the hierarchical cluster analysis. The results in session-to-session transfer and subject-transfer decoding are very significant for BCI studies as they reduced the data collection time, calibration phase, and

visual fatigue [73]–[77]. We also demonstrated a successful feasibility of online control of the proposed method with a short-time window size.

In this paper, we employed the CCA-based simple classification method because it does not need any training dataset for classification. To further improve the performance, it is necessary to incorporate more advanced SSVEP classification methods (e.g., convolutional neural network [22], correlated component analysis [23], and task-related component analysis [78]). Further, to enhance the practicality in real life, the sampled sinusoidal stimulation method [79] can be used to generate a number of visual stimuli. So far, this paper has tested the performance of ear-EEG in an SSVEP paradigm. Note that the proposed method is also applicable to other types of BCI paradigms (such as alpha attenuation, auditory onset, and P300) with the corresponding classification method.

ACKNOWLEDGMENT

The authors would like to thank K.-R. Müller and C. Guan for valuable discussions and comments. They would also like to thank K. B. Mikkelsen for providing his code on the mutual information-based prediction method.

REFERENCES

[1] M. A. Lebedev and M. A. L. Nicolelis, "Brain–machine interfaces: Past, present and future," *Trends Neurosci.*, vol. 29, no. 9, pp. 536–546, Sep. 2006.

[2] J. R. Wolpaw *et al.*, "Brain–computer interface technology: A review of the first international meeting," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 2, pp. 164–173, Jun. 2000.

[3] R. Liu, Y.-X. Wang, and L. Zhang, "An FDES-based shared control method for asynchronous brain-actuated robot," *IEEE Trans. Cybern.*, vol. 46, no. 6, pp. 1452–1462, Jun. 2016.

[4] T. Carlson and J. D. R. Millan, "Brain-controlled wheelchairs: A robotic architecture," *IEEE Robot. Autom. Mag.*, vol. 20, no. 1, pp. 65–73, Mar. 2013.

[5] K.-T. Kim, H.-I. Suk, and S.-W. Lee, "Commanding a brain-controlled wheelchair using steady-state somatosensory evoked potentials," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 3, pp. 654–665, Mar. 2018.

[6] D.-O. Won, H.-J. Hwang, D.-M. Kim, K.-R. Müller, and S.-W. Lee, "Motion-based rapid serial visual presentation for gaze-independent brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 334–343, Feb. 2018.

[7] H.-I. Suk, C.-Y. Wee, S.-W. Lee, and D. Shen, "State-space model with deep learning for functional dynamics estimation in resting-state fMRI," *NeuroImage*, vol. 129, pp. 292–307, Apr. 2016.

[8] N. Naseer and K.-S. Hong, "fNIRS-based brain–computer interfaces: A review," *Front. Human Neurosci.*, vol. 9, p. 3, Jan. 2015.

[9] J. Mellinger *et al.*, "An MEG-based brain–computer interface (BCI)," *NeuroImage*, vol. 36, no. 3, pp. 581–593, Jul. 2007.

[10] D. Iacoviello, A. Petracca, M. Spezialetti, and G. Placidi, "A classification algorithm for electroencephalography signals by self-induced emotional stimuli," *IEEE Trans. Cybern.*, vol. 46, no. 12, pp. 3171–3180, Dec. 2016.

[11] S.-Y. Dong, B.-K. Kim, and S.-Y. Lee, "EEG-based classification of implicit intention during self-relevant sentence reading," *IEEE Trans. Cybern.*, vol. 46, no. 11, pp. 2535–2542, Nov. 2016.

[12] L. Xie, Z. Deng, P. Xu, K.-S. Choi, and S. Wang, "Generalized hidden-mapping transductive transfer learning for recognition of epileptic electroencephalogram signals," *IEEE Trans. Cybern.*, vol. 49, no. 6, pp. 2200–2214. Jun. 2019.

[13] S. Lemm, B. Blankertz, T. Dickhaus, and K.-R. Müller, "Introduction to machine learning for brain imaging," *NeuroImage*, vol. 56, no. 2, pp. 387–399, May 2011.

[14] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain–computer communication," *Proc. IEEE*, vol. 89, no. 7, pp. 1123–1134, Jul. 2001.

[15] K. K. Ang and C. Guan, "EEG-based strategies to detect motor imagery for control and rehabilitation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 4, pp. 392–401, Apr. 2017.

[16] Y. Zhang, C. S. Nam, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Temporally constrained sparse group spatial patterns for motor imagery BCI," *IEEE Trans. Cybern.*, vol. 49, no. 9, pp. 3322–3332, Jun. 2019.

[17] 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.

[18] I. Käthner, A. Kübler, and S. Halder, "Rapid P300 brain–computer interface communication with a head-mounted display," *Front. Neurosci.*, vol. 9, p. 207, Jun. 2015.

[19] E. A. Aydin, O. F. Bay, and I. Guler, "P300-based asynchronous brain computer interface for environmental control system," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 3, pp. 653–663, May 2018.

[20] D. Hübner, T. Verhoeven, K.-R. Müller, P. J. Kindermans, and M. Tangermann, "Unsupervised learning for brain–computer interfaces based on event-related potentials: Review and online comparison," *IEEE Comput. Intell. Mag.*, vol. 13, no. 2, pp. 66–77, May 2018.

[21] N.-S. Kwak, K.-R. Müller, and S.-W. Lee, "A lower limb exoskeleton control system based on steady state visual evoked potentials," *J. Neural Eng.*, vol. 12, no. 5, Aug. 2015, Art. no. 056009.

[22] N.-S. Kwak, K.-R. Müller, and S.-W. Lee, "A convolutional neural network for steady state visual evoked potential classification under ambulatory environment," *PLoS ONE*, vol. 12, no. 2, Feb. 2017, Art. no. e0172578.

[23] Y. Zhang *et al.*, "Correlated component analysis for enhancing the performance of SSVEP-based brain–computer interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 5, pp. 948–956, May 2018.

[24] B. Van Dun, J. Wouters, and M. Moonen, "Improving auditory steady-state response detection using independent component analysis on multichannel EEG data," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 7, pp. 1220–1230, Jul. 2007.

[25] D.-W. Kim, J.-C. Lee, Y.-M. Park, I.-Y. Kim, and C.-H. Im, "Auditory brain–computer interfaces (BCIs) and their practical applications," *Biomed. Eng. Lett.*, vol. 2, no. 1, pp. 13–17, Mar. 2012.

[26] C. Sannelli, C. Vidaurre, K.-R. Müller, and B. Blankertz, "Ensembles of adaptive spatial filters increase BCI performance: An online evaluation," *J. Neural Eng.*, vol. 13, no. 4, May 2016, Art. no. 046003.

[27] J. Andreu-Perez, F. Cao, H. Hagras, and G.-Z. Yang, "A self-adaptive online brain–machine interface of a humanoid robot through a general type-2 fuzzy inference system," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 1, pp. 101–116, Feb. 2018.

[28] J. Minguillon, M. A. Lopez-Gordo, and F. Pelayo, "Trends in EEG-BCI for daily-life: Requirements for artifact removal," *Biomed. Signal Process. Control*, vol. 31, pp. 407–418, Jan. 2017.

[29] 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, Apr. 2017, Art. no. e0175856.

[30] A. von Lühmann, H. Wabnitz, T. Sander, and K.-R. Müller, "M3BA: A mobile, modular, multimodal biosignal acquisition architecture for miniaturized EEG-NIRS-based hybrid BCI and monitoring," *IEEE Trans. Biomed. Eng.* vol. 64, no. 6, pp. 1199–1210, Jun. 2017.

[31] C.-S. Wei, Y.-T. Wang, C.-T. Lin, and T.-P. Jung, "Toward drowsiness detection using non-hair-bearing EEG-based brain–computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 400–406, Feb. 2018.

[32] Y.-T. Wang, M. Nakanishi, Y. Wang, C.-S. Wei, C.-K. Cheng, and T.-P. Jung, "An online brain–computer interface based on SSVEPs measured from non-hair-bearing areas," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 1, pp. 14–21, Jan. 2017.

[33] S. Debener, R. Emkes, M. De Vos, and M. Bleichner, "Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear," *Sci. Rep.*, vol. 5, no. 1, Nov. 2015, Art. no. 16743.

[34] D. Looney *et al.*, "The in-the-ear recording concept: User-centered and wearable brain monitoring," *IEEE Pulse*, vol. 3, no. 6, pp. 32–42, Nov./Dec. 2012.

[35] B. Mirkovic, M. G. Bleichner, M. De Vos, and S. Debener, "Target speaker detection with concealed EEG around the ear," *Front. Neurosci.*, vol. 10, p. 349, Jul. 2016.

[36] C. B. Christensen, J. M. Harte, T. Lunner, and P. Kidmose, "Ear-EEG-based objective hearing threshold estimation evaluated on normal hearing subjects," . *IEEE Trans. Biomed. Eng.*, vol. 65, no. 5, pp. 1026–1034, May 2018.

[37] A. Nguyen *et al.*, "In-ear biosignal recording system: A wearable for automatic whole-night sleep staging," in *Proc. Workshop Wearable Syst. Appl.*, Singapore, 2016, pp. 19–24.

[38] V. Goverdovsky *et al.*, "Hearables: Multimodal physiological in-ear sensing," *Sci. Rep.*, vol. 7, no. 1, p. 6948, Jul. 2017.

[39] T. Nakamura, V. Goverdovsky, and D. P. Mandic, "In-ear EEG biometrics for feasible and readily collectable real-world person authentication," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 3, pp. 648–661, Mar. 2018.

[40] Y. Gu *et al.*, "Comparison between scalp EEG and behind-the-ear EEG for development of a wearable seizure detection system for patients with focal epilepsy," *Sensors*, vol. 18, no. 1, p. 29, Dec. 2017.

[41] D. Looney, P. Kidmose, M. J. Morrell, and D. P. Mandic, "Ear-EEG: Continuous brain monitoring," in *Brain–Computer Interface Research: A State-of-the-Art Summary 3*, C. Guger, T. Vaughan, and B. Allison, Eds. Cham, Switzerland: Springer, 2014, pp. 63–71. doi: 10.1007/978-3-319-09979-8_6.

[42] P. Kidmose, D. Looney, M. Ungsturp, and D. P. Mandic, "A study of evoked potentials from ear-EEG," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2824–2830, Oct. 2013.

[43] J.-H. Lee, S.-M. Lee, H.-J. Byeon, J.-S. Hong, K.-S. Park, and S.-H. Lee, "CNT/PDMS-based canal-typed ear electrodes for inconspicuous EEG recording," *J. Neural Eng.*, vol. 11, no. 4, Jun. 2014, Art. no. 046014.

[44] M. G. Bleichner, B. Mirkovic, and S. Debener, "Identifying auditory attention with ear-EEG: cEEGrid versus high-density cap-EEG comparison," *J. Neural Eng.*, vol. 13, no. 6, Sep. 2016, Art. no. 066004.

[45] V. Goverdovsky, D. Looney, P. Kidmose, and D. P. Mandic, "In-ear EEG from viscoelastic generic earpieces: Robust and unobtrusive 24/7 monitoring," *IEEE Sensors J.*, vol. 16, no. 1, pp. 271–277, Jan. 2016.

[46] S. L. Kappel and P. Kidmose, "High-density ear-EEG," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.(EMBC)*, Seogwipo, South Korea, 2017, pp. 2394–2397.

[47] S. L. Kappel, M. L. Rank, H. O. Toft, M. Andersen, and P. Kidmose, "Dry-contact electrode ear-EEG," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 1, pp. 150–158. Jan. 2019.

[48] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki, "EmotionMeter: A multimodal framework for recognizing human emotions," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 1110–1122. Mar. 2019.

[49] D. Looney, P. Kidmose, and D. P. Mandic, "Ear-EEG: User-centered and wearable BCI," in *Brain–Computer Interface Research: A State-of-the-Art Summary 2*, C. Guger, B. Allison, and E. C. Leuthardt, Eds. Heidelberg, Germany: Springer, 2014, pp. 41–50. doi: 10.1007/978-3-642-54707-2_5.

[50] Y. T. Wang, Y. Wang, C. K. Cheng, and T. P. Jung, "Measuring steady-state visual evoked potentials from non-hair-bearing areas," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.(EMBC)*, San Diego, CA, USA, 2012, pp. 1806–1809.

[51] C. S. Herrmann "Human EEG responses to 1–100 Hz flicker: Resonance phenomena in visual cortex and their potential correlation to cognitive phenomena," *Exp. Brain Res.*, vol. 137, nos. 3–4, pp. 346–353, Apr. 2001.

[52] K. B. Mikkelsen, P. Kidmose, and L. K. Hansen, "On the keyhole hypothesis: High mutual information between ear and scalp EEG," *Front. Human Neurosci.*, vol. 11, p. 341, Jul. 2017.

[53] L. Paninski, "Estimation of entropy and mutual information," *Neural Comput.*, vol. 15, no. 6, pp. 1191–1253, Jun. 2003.

[54] D. Wu, J.-T. King, C.-H. Chuang, C.-T. Lin, and T.-P. Jung, "Spatial filtering for EEG-based regression problems in brain–computer interface (BCI)," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 2, pp. 771–781, Apr. 2018.

[55] D. Wu, V. J. Lawhern, S. Gordon, B. J. Lance, and C.-T. Lin, "Driver drowsiness estimation from EEG signals using online weighted adaptation regularization for regression (OwARR)," *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 6, pp. 1522–1535, Dec. 2017.

[56] J.-H. Kim, F. Bießmann, and S.-W. Lee, "Decoding three-dimensional trajectory of executed and imagined arm movements from electroencephalogram signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 5, pp. 867–876, Sep. 2015.

[57] C. J. Stam, "Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field," *Clin. Neurophysiol.*, vol. 116, no. 10, pp. 2266–2301, Oct. 2005.

[58] Y. A. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller, "Efficient backprop," in *Neural Networks: Tricks of the Trade*. Heidelberg, Germany: Springer, 2012, pp. 9–48. doi: 10.1007/978-3-642-35289-8_3.

[59] R. Ramakrishnan, P. O. Dral, M. Rupp, and O. A. von Lilienfeld, "Big data meets quantum chemistry approximations: The Δ-machine learning approach," *J. Chem. Theory Comput.*, vol. 11, no. 5, pp. 2087–2096, Apr. 2015.

[60] Z. Lin, C. Zhang, W. Wu, and X. Gao, "Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 6, pp. 1172–1176, Jun. 2007.

[61] G. Bin, X. Gao, Z. Yan, B. Hong, and S. Gao, "An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method," *J. Neural Eng.*, vol. 6, no. 4, Aug. 2009, Art. no. 046002.

[62] Y. U. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, "Frequency recognition in SSVEP-based BCI using multiset canonical correlation analysis," *Int. J. Neural Syst.*, vol. 24, no. 4, Jun. 2014, Art. no. 1450013.

[63] L. S. Aiken, S. G. West, and S. C. Pitts, "Multiple linear regression," in *Comprehensive Handbook of Psychology: Research Methods in Psychology*, J. Schinka and W. Velicer, Eds. New York, NY, USA: Wiley, 2003, pp. 481–507.

[64] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 42, no. 1, pp. 80–86, Feb. 2000.

[65] V. Vovk, "Kernel ridge regression," in *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik*, B. Scholkop, Z. Luo, and V. Vovk, Eds. Heidelberg, Germany: Springer, 2013, pp. 105–116. doi: 10.1007/978-3-642-41136-6_11.

[66] B. Schölkopf, A. Smola, and K.-R. Müller, "Kernel principal component analysis," in *Proc. Int. Conf. Artif. Neural Netw.*, Lausanne, Switzerland, 1997, pp. 583–588.

[67] T.-P. Jung, S. Making, M. Stensmo, and T. J. Sejnowski, "Estimating alertness from the EEG power spectrum," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 1, pp. 60–69, Jan. 1997.

[68] C.-S. Wei, Y.-P. Lin, Y.-T. Wang, C.-T. Lin, and T.-P. Jung, "A subject-transfer framework for obviating inter- and intra-subject variability in EEG-based drowsiness detection," *NeuroImage*, vol. 174, pp. 407–419, Jul. 2018.

[69] E. Yin, Z. Zhou, J. Jiang, Y. Yu, and D. Hu, "A dynamically optimized SSVEP brain–computer interface (BCI) speller," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1447–1456, Jun. 2015.

[70] M. Wang, R. Li, R. Zhang, G. Li, and D. Zhang, "A wearable SSVEP-based BCI system for quadcopter control using head-mounted device," *IEEE Access*, vol. 6, pp. 26789–26798, 2018.

[71] L. Acqualagna *et al.*, "EEG-based classification of video quality perception using steady state visual evoked potentials (SSVEPs)," *J. Neural Eng.*, vol. 12, no. 2, Mar. 2015, Art. no. 026012.

[72] M. Dannhauer, B. Lanfer, C. H. Wolters, and T. R. Knösche, "Modeling of the human skull in EEG source analysis," *Human Brain Mapping*, vol. 32, no. 9, pp. 1383–1399, Sep. 2011.

[73] V. Jayaram, M. Alamgir, Y. Altun, B. Scholkopf, and M. Grosse-Wentrup, "Transfer learning in brain–computer interfaces," *IEEE Comput. Intell. Mag.*, vol. 11, no. 1, pp. 20–31, Feb. 2016.

[74] M. Nakanishi, Y. Wang, and T.-P. Jung, "Session-to-session transfer in detecting steady-state visual evoked potentials with individual training data," in *Foundations of Augmented Cognition: Neuroergonomics and Operational Neuroscience*, D. Schmorrow and C. Fidopiastis, Eds. Cham, Switzerland: Springer, 2016, pp. 253–260. doi: 10.1007/978-3-319-39955-3_24.

[75] C.-S. Wei, M. Nakanishi, K.-J. Chiang, and T.-P. Jung, "Exploring human variability in steady-state visual evoked potentials," in *Proc. IEEE Int. Conf. Syst. Man Cybern. (SMC)*, Miyazaki, Japan, 2018, pp. 474–479.

[76] S. Fazli, F. Popescu, M. Danóczy, B. Blankertz, K.-R. Müller, and C. Grozea, "Subject-independent mental state classification in single trials," *Neural Netw.*, vol. 22, no. 9, pp. 1305–1312, Nov. 2009.

[77] M. Krauledat, M. Tangermann, B. Blankertz, and K.-R. Müller, "Towards zero training for brain–computer interfacing," *PLoS ONE*, vol. 3, no. 8, Aug. 2008, Art. no. e2967.

[78] M. Nakanishi, Y. Wang, X. Chen, Y.-T. Wang, X. Gao, and T.-P. Jung, "Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 1, pp. 104–112, Jan. 2018.

[79] X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao, "Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface," *J. Neural Eng.*, vol. 12, no. 4, Aug. 2015, Art. no. 046008.

No-Sang Kwak received the B.S. degree in electronic engineering from the Seoul National University of Science and Technology, Seoul, South Korea, in 2013. He is currently pursuing the Ph.D. degree in brain cognitive engineering with Korea University, Seoul.

His current research interests include machine learning, deep learning, and brain–computer interface.

Seong-Whan Lee (F'10) received the B.S. degree in computer science and statistics from Seoul National University, Seoul, South Korea, in 1984 and the M.S. and Ph.D. degrees in computer science from the Korea Advanced Institute of Science and Technology, Seoul, in 1986 and 1989, respectively.

He is currently the Hyundai-Kia Motor Chair Professor and the Head of the Department of Artificial Intelligence, Korea University, Seoul. His current research interests include pattern recognition, artificial intelligence, and brain engineering.

Dr. Lee is a fellow of IAPR and the Korea Academy of Science and Technology.