Extraction of SSVEPs-based Inherent Fuzzy Entropy Using a Wearable Headband EEG in Migraine Patients

Zehong Cao, *Member, IEEE*, Chin-Teng Lin*, *Fellow, IEEE*, Kuan-Lin Lai, Li-Wei Ko, *Member, IEEE*, Jung-Tai King, Kwong-Kum Liao, Jong-Ling Fuh, and Shuu-Jiun Wang*

Abstract - **Inherent fuzzy entropy is an objective measurement of electroencephalography (EEG) complexity, reflecting the robustness of brain systems. In this study, we present a novel application of multi-scale relative inherent fuzzy entropy using repetitive steady-state visual evoked potentials (SSVEPs) to investigate EEG complexity change between two migraine phases, i.e. inter-ictal (baseline) and pre-ictal (before migraine attacks) phases. We used a wearable headband EEG device with O1, Oz, O2 and Fpz electrodes to collect EEG signals from 80 participants (40 migraine patients and 40 healthy controls [HCs]) under the following two conditions: during resting state and SSVEPs with five 15-Hz photic stimuli. We found a significant enhancement in occipital EEG entropy with increasing stimulus times in both HCs and patients in the inter-ictal phase but a reverse trend in patients in the pre-ictal phase. In the 1st SSVEP, occipital EEG entropy of the HCs was significantly lower than that of patents in the pre-ictal phase (FDR-adjusted p < 0.05). Regarding the transitional variance of EEG entropy between the 1st and 5th SSVEPs, patients in the pre-ictal phase exhibited significantly lower values than patients in the inter-ictal phase (FDR-adjusted p < 0.05). Furthermore, in the classification model, the AdaBoost ensemble learning showed an accuracy of 81±6% and AUC of 0.87 for classifying inter-ictal and pre-ictal phases. In contrast, there were no differences in EEG entropy among groups or sessions by using other competing entropy models, including approximate entropy, sample entropy and fuzzy entropy on the same dataset. In conclusion, inherent fuzzy entropy offers novel applications in visual stimulus environments and may have the potential to provide a pre-ictal alert to migraine patients.**

Index Terms - **Migraine, SSVEP, EEG, Inherent Fuzzy Entropy**

I. INTRODUCTION

Migraine is a type of neurovascular headache that presents a severe throbbing head pain and is accompanied by nausea, vomiting or extreme sensitivity to light and sound [1]. Episodic migraine is considered a recurrent headache with a cycle that includes inter-ictal, pre-ictal, ictal and post-ictal phases [2]. The ictal phase is the period during which migraine patients suffer from headache; the pre-ictal phase is defined as 72 hours before the ictal phase, which is preceded by the inter-ictal phase.

The activation and sensitization of brain activities, which have powerful potential applications, can be characterized by the visual stimulus environment [3]. In particular, steady-state visual evoked potentials (SSVEPs) are responses to photic stimuli at a multiple of or equal to the frequency of the stimuli. The visual system of the brain is intimately connected with the environment via the eyes, which contain light receptors on the retina. These receptors send electrical signals to the brain, which can be detected by cortical electroencephalography (EEG). These messages are generally sent to the occipital region [4], which causes cortical activation directing to specific frequency stimulations [5]. Migraine patients are more vulnerable to visual stimuli, and previous studies have reported that visual stimuli are effective for examining the dishabituation in the inter-ictal or pre-ictal phases [6-11]. In previous SSVEP studies, it has been shown that the cortical responses would be enhanced in patients with migraine as compared to those of control subjects, in a broad range of stimulation frequency, supporting the concept of central hyper-responsiveness in patients with migraine. Initially, stimulation frequencies at median-to-high range (i.e., 15 - 30 Hz) were demonstrated to have such differentiating property [12, 13]. Lately, it was realized that stimulations at lower frequencies (i.e., 3 - 6 Hz)

were also able to elicit such differences [14]. Considering overt visual stimulation could evoke migraine attacks, or even provoke seizures [15], we thus decided to use only one stimulation frequency (i.e., 15 Hz), which was well located in the middle of the frequency distribution, to avoid the possible adverse effects.

To efficiently extract features from SSVEP-based models, multivariate linear regression [16] and multi-set/-layer/-way canonical correlation analysis [17-19] were developed and focused on frequency recognition. In addition to the above measurements, recently developed entropy analysis approaches in the temporal domain have helped us to understand brain dynamics and allowed us to assess how complexity provides information about a wide range of physiological systems [20]. Entropy is generally an objective measure of how the complexity of physiological signals represents the robustness of brain systems [21]. Different entropy analysis approaches, such as approximate entropy (ApEn) [22] and sample entropy (SampEn) [23] have been developed to measure complex signals.

Fuzziness, featured by uncertainty, can avoid a sharp distinction of the boundary of a set [24]. For example, the fuzzy entropy (FuzzEn) algorithm [25, 26] incorporates a fuzzy membership function instead of the Heaviside function to assess the degree of similarity between two vectors' shapes, which can effectively overcome the deficiency in existing subspace filtering techniques [27]. Integrated brain systems are often multi-scaled and can interact with fast or slow processes, depending on the scale of the bio-signal of interest. Thus, we recently developed a multi-scale inherent fuzzy entropy (Inherent FuzzEn) algorithm [28] that has the robustness to operate under noise, nonlinear and non-stationary signals and is capable of operating on EEG signals across a range of temporal (time) scales. Compared to ApEn, SampEn or FuzzEn, the Inherent FuzzEn algorithm shows stable complexity and the smallest root mean square deviation in the resting-state condition. Therefore, we applied the Inherent FuzzEn algorithm to explore resting-state EEG complexity before migraine attacks [29], which showed that the EEG complexity of patients in the pre-ictal phase was significantly higher than that of patients in the inter-ictal phase in the prefrontal area.

To the best of our knowledge, high and low complexity represent healthy (robust) and diseased (vulnerable) brain systems [20, 29]. We hypothesize that a healthy brain system demonstrates strong robustness after repetitive visual stimulation (linked to habituation) [30] but that migraine patients may exhibit less robustness of the brain system. However, existing studies have not investigated the brain complexity in migraine patients as well as healthy controls (HCs) using entropy analysis approaches in repetitive visual stimulus sessions. Thus, we aimed to 1) understand the effects of repetitive SSVEPs in migraine patients and HCs by measuring inherent fuzzy entropy; 2) classify inter-ictal and pre-ictal migraine phases using input features of inherent fuzzy entropy; 3) compare our method to other entropy algorithms (ApEn, SampEn and FuzzEn). Our study represents a novel application of inherent fuzzy entropy in a visual stimulus environment for migraine patients, and the extracted features have the potential to provide pre-ictal alerts to migraine patients.

This paper is organized as follows. First, a relative inherent fuzzy entropy algorithm that minimizes individual differences is presented in Section II. Section III introduces the experimental procedures, including the recruitment of participants, the wearable EEG device, the experimental paradigm, the data processing classification models and the statistical analysis. In Section IV, the application performances of the relative inherent fuzzy entropy algorithm are assessed regarding migraine patients and HCs. Section V discusses the experimental results. Section VI provides the conclusion.

II. EEG COMPLEXITY: RELATIVE MULTI-SCALE INHERENT FUZZY ENTROPY

EEG complexity is measured by the relative multi-scale inherent fuzzy entropy algorithm [28], which is divided into three parts: A. a de-trending process; B. a multi-scale procedure; and C. fuzzy entropy assessment. In this study, to minimize individual differences, we estimated the differences in multi-scale inherent fuzzy entropy between the baseline (resting) and stimulus sessions (refer to part D), termed the relative multi-scale inherent fuzzy entropy (RE). As shown in Fig. 1, we present an overview of our proposed relative multi-scale inherent fuzzy entropy.

A. De-trending process to extract inherent modes

We applied empirical mode decomposition to decompose the raw EEG signal $s(t)$ into several intrinsic mode functions and reconstructed the signal $\hat{s}(t)$.
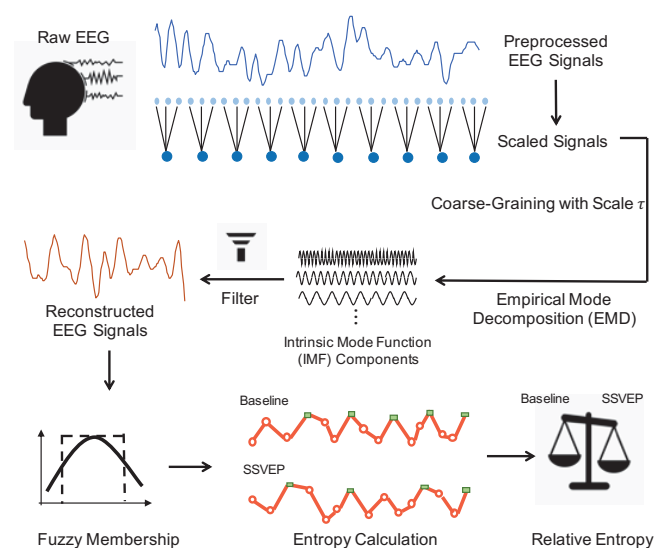

Figure 1 An overview of relative multi-scale inherent fuzzy entropy

In the initial step, extrema of the signal $s(t)$ are found, corresponding to E_{minima} and E_{maxima}. Then, the regions between E_{minima} and E_{maxima} are interpolated, yielding an envelope with $en_{min}(t)$ and $en_{max}(t)$.

First, we compute the mean:

$$M(t) = (en_{min}(t) + en_{max}(t))/2 \tag{1}$$

Second, we extract the candidate of inherent functions:

$$Ca(t) = s(t) - M(t) \tag{2}$$

Third, we confirm $Ca(t)$, belonging to an intrinsic mode function. If $Ca(t)$ satisfies the constraint conditions, $Ca(t)$ is saved, and the residue is computed:

$$res(t) = s(t) - \sum_{i=1}^{t} Ca(t) \tag{3}$$

Next, we solve $t = t + 1$ and treat $res(t + 1)$ as input data. Otherwise, we treat $Ca(t + 1)$ as input data. Iterations are performed on the residual $res(t)$ and continued until the final residue r satisfies the stopping criterion.

Finally, the components of intrinsic mode functions surviving high trends are automatically removed by a trend filter. The signal $\hat{s}(t)$ is reconstructed by the cumulative sum of the remaining intrinsic mode functions:

$$\hat{s}(t) = \sum_{i=n}^{i=m} Ca(t) \tag{4}$$

The parameter i is the order number of the components from the intrinsic mode functions, and parameters m and n are the upper and lower boundaries of the selected components, respectively.

B. Multi-scale procedure

The initial step is to normalize the EEG signal using the Z-score measurement. The EEG signal $\hat{s}(t)$ subtracts the mean prior to dividing by the standard deviation. The normalized EEG signal is marked as $x(t)$. Afterwards, the multi-scale procedure involves coarse-graining the signals into different time scales.

For a given time series, multiple coarse-grained time series are constructed by averaging the data points within non-overlapping windows of increasing lengths, and the τ element of the coarse-grained time series $\hat{x}_j^{(\tau)}$ is expressed as:

$$\hat{x}_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i \tag{5}$$

where τ represents the scale factor, and $1 \leq j \leq N/\tau$.

C. Fuzzy entropy assessment

First, considering the N sample time series $\{\hat{x}(i): 1 \leq i \leq N\}$, given m, n, and r, a vector set sequence $\{X_i^m, i = 1, ..., N - m + 1\}$ is calculated, and the baseline is removed:

$$X_i^m = \begin{Bmatrix} \hat{x}(i), \hat{x}(i + 1) ..., \\ \hat{x}(i + m - 1) \end{Bmatrix} - m^{-1} \sum_{j=0}^{m-1} \hat{x}(i + j) \tag{6}$$

where $1 \leq i \leq N - m + 1$, and X_i^m presents m consecutive u values, beginning with the ith point.

Second, when judging whether a vector X_j is similar to X_i within a tolerance r, in sample entropy, the similarity degree can be expressed simply as follows:

$$C_{ij}(d_{ij}, r) = \begin{cases} 1, & if\ d_{ij} < r \\ 0, & if\ d_{ij} > r \end{cases}$$
$$\tag{7}$$

where d_{ij} represents the distance between the two vectors X_j and X_i, and function C_{ij} is Heaviside function serving as a two-state classifier.

Since the boundary of Heaviside function is rigid, in fuzzy entropy, given a vector X_i^m, the similarity degree D_{ij}^m between X_i^m and X_j^m is defined by the fuzzy membership function:

$$D_{ij}^m = fu(d_{ij}^m, n, r) = exp\left(-\frac{(d_{ij}^m)^n}{r}\right) \tag{8}$$

where the fuzzy membership function fu is an exponential function that is more appropriate for processing physiological signals in term of robustness to noise and independence on the data length [25, 31, 32].

The parameter m is the length of the sequences to be compared to other entropy algorithms. The other two parameters, r and n, determine the width and the gradient of the boundary of the fuzzy membership function, respectively.

Then, the function φ^m is constructed. Similarly, for $m + 1$, the above steps are repeated and denoted $\varphi^{m+1}(n, r)$.

$$\varphi^m(n, r) = (N - m)^{-1} \sum_{i=1}^{N-m}((N - m - 1)^{-1} \sum_{j=1, j \neq i}^{N-m} D_{ij}^m)$$
$$\tag{9}$$

Finally, the $entropy\ (m, n, r, N)$ parameter of the sequence $\{\hat{x}(i): 1 \leq i \leq N\}$ is defined as the negative natural logarithm of the deviation of φ^m from φ^{m+1}:

$$entropy(m, n, r, N) = \ln \varphi^m(n, r) - \ln \varphi^{m+1}(n, r) \tag{10}$$

Of note, the measurement of Eq. (10) is consistent with the previously developed approximate entropy, sample entropy, and fuzzy entropy. The entropy measures the likelihood that runs of the patterns that are close to the observations remain close on the next incremental comparisons, which can quantify physiological time-series complexity. Please see below the proof to verify that the measurement in Eq. (10) represents an entropy.

Proof Eq. (10):

Let B_{ij} be the number of vectors X_i^m within r of X_j^m and Let A_{ij} be the number of vectors X_i^{m+1} within r of X_j^{m+1}.

Define the function $D_{ij}^m(n, r) = (B_{ij})/(N - m)$ and $D_{ij}^{m+1}(n, r) = (A_{ij})/(N - m + 1)$.

$D_{ij}^m(n,r)$ is the probability that any vector X_i^m is within r of X_j^m by the fuzzy membership function.

$D_{ij}^{m+1}(n,r)$ is the probability that any vector X_i^{m+1} is within r of X_j^{m+1} by the fuzzy membership function.

Then, based on the Grassberger and Procaccia's work [33], define the function:

$$\varphi^m(n,r) = (N-m)^{-1}\sum_{i=1}^{N-m} In\,[D_{ij}^m(r)]$$

which is the average of natural logarithms of functions $C_i^m(n,r)$

Similarly,

$$\varphi^{m+1}(n,r) = (N-m+1)^{-1}\sum_{i=1}^{N-m+1} In\,[D_{ij}^{m+1}(n,r)]$$

With the limitation as $N \to \infty$,

$$\varphi^m(n,r) = (N-m)^{-1}\sum_{i=1}^{N-m} In\,[D_{ij}^m(r)]$$

$$\varphi^{m+1}(n,r) = (N-m)^{-1}\sum_{i=1}^{N-m} In\,[D_{ij}^{m+1}(r)]$$

Based on Pincus's work [22], entropy measures the likelihood that runs of the patterns that are close to m observations remain close on the next incremental comparisons.

$$Entropy = \ln\varphi^m(n,r) - \ln\varphi^{m+1}(n,r)$$
$$= (N-m)^{-1}\{\sum_{i=1}^{N-m} In\,[D_{ij}^m(n,r) - InD_{ij}^{m+1}(n,r)]\}$$

which equals the average over i of $In\,[D_{ij}^m(n,r)/D_{ij}^{m+1}(n,r)] = In\,[B_{ij}/A_{ij}]$.

We noted that the ratio $[B_{ij}/A_{ij}]$ represents an entropy.

D. Relative inherent fuzzy entropy

To minimize individual differences, we introduced a relative multi-scale inherent fuzzy entropy (RE). First, the inherent fuzzy entropy at baseline (resting) and during stimulus sessions is calculated. The two sessions are simplified to $Entropy_{baseline}$ and $Entropy_{SSVEP}$, respectively.

Next, we calculated the variation in the multi-scale inherent fuzzy entropy between the baseline (resting) condition and the stimulus condition during which SSVEPs were induced in five stimulus trials. This function is termed the relative inherent fuzzy entropy, which is expressed as:

$$RE_k = Entropy_{SSVEP(k)} - Entropy_{baseline} \qquad (11)$$

where k is the stimulus time.

III. Experimental Procedures and Methods

A. Participants

Forty outpatients with migraine without aura (F:M = 30:10, mean age: 38.1 ± 8.2) were recruited from the Headache Clinic of Taipei Veterans General Hospital and were asked to keep a headache diary to determine migraine phases on a daily basis. All enrolled patients fulfilled the diagnostic criteria of the International Classification of Headache Disorders 2nd edition (ICHD-II) and had a migraine frequency ranging from 1 to 6 days per month. Forty age- and sex-matched HCs (F:M = 32:8, mean age: 36.1 ± 9.8) were recruited from hospital colleagues or their relatives or friends. The individuals who served as HCs did not have a past medical history or a family history of migraine. As we used the same dataset in another recent study [29], the comparisons of demographics, headache profile, and psychological characteristics between HCs and patients are summarized in that report.

As shown in Fig. 2, the days on which the EEG examinations were performed were classified into one of four migraine phases (inter-ictal, pre-ictal, ictal, or post-ictal) based on the headache diary. The ictal phase was coded when the patient was suffering from a migraine attack on the day of the EEG study. Based on previous criteria [34, 35], the pre-ictal and post-ictal phases were coded on the day of the EEG study if the patient was within 72 hours before or after an ictal phase, respectively. The inter-ictal phase was coded if the patient had not had a migraine attack within 72 hours before or after the EEG examination. Only EEG data collected during the inter-ictal and pre-ictal phases were selected for analysis in this study, as these periods are significant phases before migraine attacks and potential to detect the transition of migraine phases.

All the participants had normal vision and no systemic diseases, connective tissue disorders, neurological, psychiatric disorders, or other painful conditions according to their self-reports. None of our migraine patients received preventive treatment, and they were asked not to take any analgesics within two days before the EEG recording. The Institutional Review Board of Taipei Veterans General Hospital approved this study. Informed consent was obtained from all participants before they entered in the study.
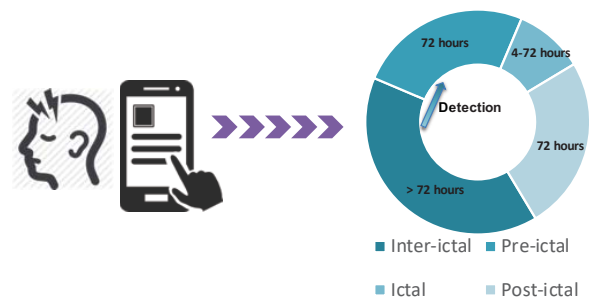

Figure 2 Headache diary and migraine cycle.

B. Wearable EEG device

EEG signals were recorded at a sampling rate of 500 Hz by a "Mindo" EEG device (Brain Rhythm Inc., Zhubei District, Hsinchu, Taiwan), which is a wearable headband EEG device with dry sensors [36]. Each dry-contact electrode was designed to include a probe head, a plunger, a spring, and a barrel. The probes were inserted into a flexible substrate via an established injection molding procedure using a one-time forming process. These dry electrodes are more convenient for measuring EEG signals than conventional wet electrodes and are preferred because they avoid the need to use conductive gel and extensive skin preparation procedures while achieving a signal quality comparable to that of wet electrodes. In this study, as shown in Fig. 3-A, four dry-contact electrodes (Fpz, O1, Oz, and O2) were placed according to the extended International 10–20 system, and two extra channels (A1 and A2) were used as reference channels. The collected EEG signals were transmitted by Bluetooth to a personal computer.

C. Experimental paradigm

The EEG experiment was performed in a static room at Taipei Veterans General Hospital, Taiwan. To avoid light source interference, we turned the fluorescent lamps off during the experimental procedure. As shown in Fig. 3-B, the illumination of monitor (Viewsonic V3D231) was 90 lux s/pulse at 20 cm in front of the participants' eyes, and the participants were asked to put their chin on the shelf. Additionally, the screen offers the functions to inspect the raw EEG signals and stimulate repetitive visual stimulus flicks in the form of alternating graphical patterns.

As shown in Fig. 3-C and Fig. 3-D, the experiment consisted of the following two sessions: a resting session and a visual stimulus session with five stimulus trials. The resting session comprised three epochs, during which the eyes were open for 1 min and closed for 1 min. To prevent the migraine patients from receiving excessive visual stimuli, we administered five trials

Figure 3 Experimental paradigm. (A) Wearable headband EEG device; (B) experimental environment and setting; (C) experiment during resting condition; (D) experiment during SSVEP condition.

of 10-second, 15-Hz stimuli to evoke SSVEPs with 10-second intervals between stimuli in the eyes-closed condition. To evaluate the EEG complexity changes related to visual stimulation, resting EEG recording during eye-closed condition was served as the baseline [37] and was then extracted to calculate the relative multi-scale inherent fuzzy entropy.

Of note, patients and HCs participated in five sessions of identical resting-state and SSVEP EEG examinations over 3–7 months, and each examination was separated by an interval of 2-8 weeks. This protocol aimed to acquire EEG data for patients who experienced different phases of migraine, particularly the inter-ictal and pre-ictal phases.

D. EEG processing

All the EEG data were analyzed using the EEGLAB toolbox (http://sccn.ucsd.edu/eeglab/) with MATLAB software (Mathworks, Inc.), an open source toolbox for electrophysiological signal processing. The EEG signals were recorded from the O1, Oz, O2 and Fpz electrodes for the resting and SSVEP epochs during the eyes-closed condition.

1. Pre-processing

The original EEG was reviewed by experienced EEG specialists. EEG activity was sampled at 500 Hz, down-sampled to 250 Hz, and then filtered through 1-Hz high-pass and 30-Hz low-pass finite impulse-response filters. Segments contaminated with non-physiological artifacts, including movement artifacts, electrode detachment, sweating artifacts, or 60-Hz noise were marked and discarded.

2. CCA-based stimulus signal enhancement

As occipital area is sensitive with SSVEP, prior studies demonstrated the power spectrum has a distinct peak at certain stimulus frequencies (e.g., 15 Hz). In this study, we expect that CCA-based method can enhance the signal quality, such as a more apparent amplitude peak at certain stimulus frequencies. Several studies have applied canonical correlation analysis (CCA) to improve signal quality [38-40], and we also have reported that the quality of EEG signals can be enhanced by CCA-based algorithms [41]. In this study, the filtered EEG signal $s(x)$ in stimulus sessions was expressed as:

$$\text{EEG data: } S(x) = \begin{bmatrix} O1 \\ Oz \\ O2 \\ Fpz \end{bmatrix} \tag{12}$$

The template sinusoidal and cosinusoidal signals associated with the flickering frequency $S(y)$ were expressed as:

$$\text{Template signal: } S(y) = \begin{bmatrix} \sin(2\pi f_1 n) \\ \cos(2\pi f_1 n) \\ \sin(2\pi f_2 n) \\ \cos(2\pi f_2 n) \end{bmatrix} \tag{13}$$
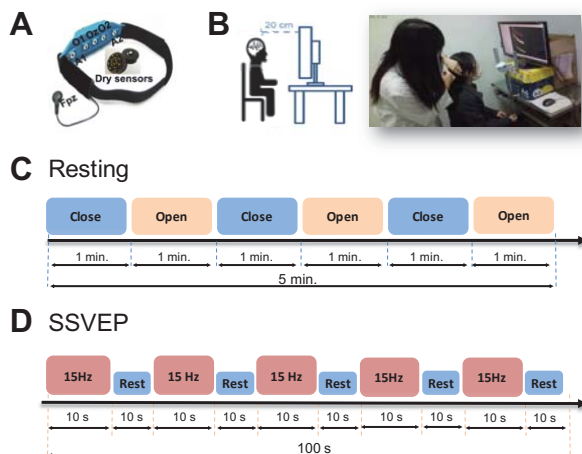
where f_1 is the frequency of the stimulation, f_2 is the second harmonic frequency to the stimulus frequency, and n is the length of the EEG signal according to the following equation:

$$n = \frac{1}{f_s} \qquad (14)$$

where f_s is the sampling rate.

The template sine and cosine signals were 15 Hz, and the template second harmonic sine and cosine signals were 30 Hz.

Then, we calculated the correlation between $S(x)$ and $S(y)$:

$$\left[C_{xx}^{-1} C_{xy} C_{yy}^{-1} C_{yx} \right] W_x = \rho^2 W_x \qquad (15)$$

where C_{xx} is the covariance matrix of $S(x)$, and C_{yy} is the covariance matrix of $S(y)$. C_{xy} and C_{yx} are the cross-covariance matrices between $S(x)$ and $S(y)$. W_x contains the eigenvectors and eigenvalues of matrices $S(x)$ and $S(y)$.

Considering the components of the eigenvector that are similar to the stimulus frequency, the first two selected components were comprised matrix A. In addition, we transposed matrix A and multiplied it by the original EEG data $S(x)$:

$$Y' = A^T S(x) \qquad (16)$$

The inverse of matrix A was also transposed and multiplied by matrix Y':

$$S(x)' = (A^T)^{-1} Y' \qquad (17)$$

The matrix $S(x)'$ represents the EEG signals after CCA-based artifact removal.

Finally, the artifact-free EEG signals were inspected again using the automatic continuous rejection function in EEGLAB.

3. Entropy estimation

The CCA-based EEG data were estimated by the relative multi-scale inherent fuzzy entropy algorithm (refer to Section II). During the multi-scale procedure, each coarse-grained time series was calculated for different temporal scales from 1 to 20. The parameters we used in this study are $m = 2$, $r = 0.15$ and $\tau = 20$. Of note, the values of the relative inherent fuzzy entropy in the resting and stimulus conditions were obtained by averaging the entropy estimates of three 1-min eyes-closed blocks and 10-s SSVEP blocks, respectively. The prefrontal entropy was calculated from the entropy values calculated from data recorded from the Fpz electrode, and the occipital entropy was calculated by averaging the entropy values corresponding to the O1, Oz, O2 electrodes. The entropy estimates for each patient were averaged over the same migraine phase if the participant had more than one EEG recordings during inter-ictal or pre-ictal phase, and the entropy estimates of each HC were averaged over the examinations.

E. Classification models

A binary classification model was developed to discriminate among migraine phases (inter-ictal vs. pre-ictal). In this study, we employed six commonly used algorithms, including linear discriminant analysis (LDA), a k-nearest neighbors classifier (kNN), multilayer perceptron (MLP), a Bayesian classifier, a support vector machine (SVM) with a linear or radial basis function (RBF) kernel [42], Gaussian processes (GP) [43],

random forest (RF) [44], and adaptive boosting (AdaBoost) [45] to classify migraine phases based on EEG entropy features with significant levels. These classification algorithms were all implemented using PRTools [46], LIBSVM [47], GPML [43] or MATLAB File Exchanges in MATLAB software.

The performances of the six algorithms were validated and compared via a 3-fold cross-validation procedure. That is, all data were randomly partitioned into 3 approximately equally-sized clusters. Two clusters were combined and used as the training data, and the remaining cluster was retained as the validation data to test the model. For example, the original sample (40 patients) was randomly partitioned into 3 equally-sized subsamples. Of the 3 subsamples, 2 subsamples (~27 patients) were used as training data, and the remaining subsample (~13 patients) was retained as the validation data for testing the model. The cross-validation process was repeated 100 times. Additionally, due to the limitation of a small dataset, the approach for optimizing hyperparameter is tuned by cross-validation (k-fold) to evaluate the model performance. Specifically, we first split train and test subsets with 3-folds, and randomized search of parameters on train subset. Then, we selected the best estimator obtained after the randomized search based on accuracy. Finally, we tested the tuned classifiers with test subset and obtained the accuracy scores.

The performance metrics covered classification accuracy, recall, precision, and F-measure [48]. Specifically, accuracy is the most intuitive performance measure, which is simply a ratio of the correctly predicted observations to the total observations. Recall (also called sensitivity) is the ratio of correctly predicted positive observations to all observations in the actual class. Precision (also called positive predictive value) is the ratio of correctly predicted positive observations to the total predicted positive observations. F-measure is the weighted average of precision and recall, which takes both false positives and false negatives into account. Additionally, the receiver operating characteristic (ROC) curves and area under the curve (AUC) were employed to evaluate the performance of various classifiers.
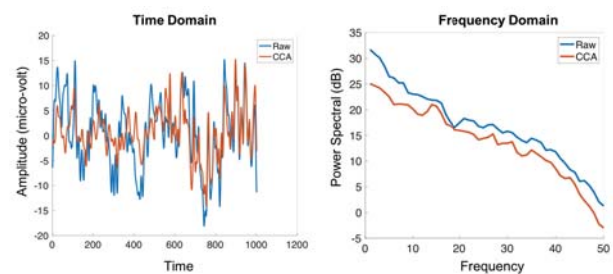

Figure 4 CCA-based signal enhancement

F. Statistical analysis

To determine the independence of different entropy variables in of repetitive visual stimulation (i.e. between the 1st and 5th stimuli), we performed paired *t*-tests for each entropy scale to compare EEG complexity within the same group (HCs, inter-ictal, and pre-ictal patients). Furthermore, paired *t*-tests were applied to compare EEG complexity between inter-ictal and pre-ictal patients, and independent *t*-tests were used to test for differences between groups (HCs vs. inter-ictal or pre-ictal patients). False discovery rate (FDR) correction was used to control for multiple comparisons. The intra-class correlation was estimated to quantify a test-retest reliability. One-way ANOVA was used to compare the performances of the classification algorithms, followed by Tukey's post hoc test to test all pairwise comparisons. All the statistical tests were two-tailed, and statistical significance was set at $p < 0.05$.

IV. RESULTS

Firstly, we compared the temporal EEG signals and corresponding power spectra before and after the CCA procedure to ensure a more apparent amplitude peak at stimulus frequency. Here, as shown in Fig. 4, we used an EEG signal with 15Hz SSVEP to compare the temporal EEG signals and corresponding power spectra before and after the CCA procedure. The left figure presented the raw (blue line) and CCA processed (red line) EEG signals in the time domain, and the right figure demonstrated the corresponding power spectra before (blue line) and after (red line) the CCA procedure. The CCA-based signal enhancement results showed a more apparent peak at stimulus frequency 15Hz and harmonic frequency 30Hz, relative to the raw EEG signals.

Furthermore, EEG complexity was measured with multi-scale relative inherent fuzzy entropy (RE), simply referred to as "entropy" in the following sections, over different time scales τ ranging from 1 to 20 in repetitive SSVEPs from prefrontal and occipital areas. With increased stimulus times, we noted monotonic enhancement in the occipital EEG entropy of HCs and patients during the inter-ictal phase, but a monotonic reduction in the occipital EEG entropy in patients during the pre-ictal phase (Fig. 5-A). However, no significant differences or trends in the EEG entropy were observed among the different groups in the prefrontal region. Therefore, in the following sections, we present the detailed findings of the occipital regions.

A. Intra-group comparisons of EEG entropy

In this section, we compared the differences in EEG entropy
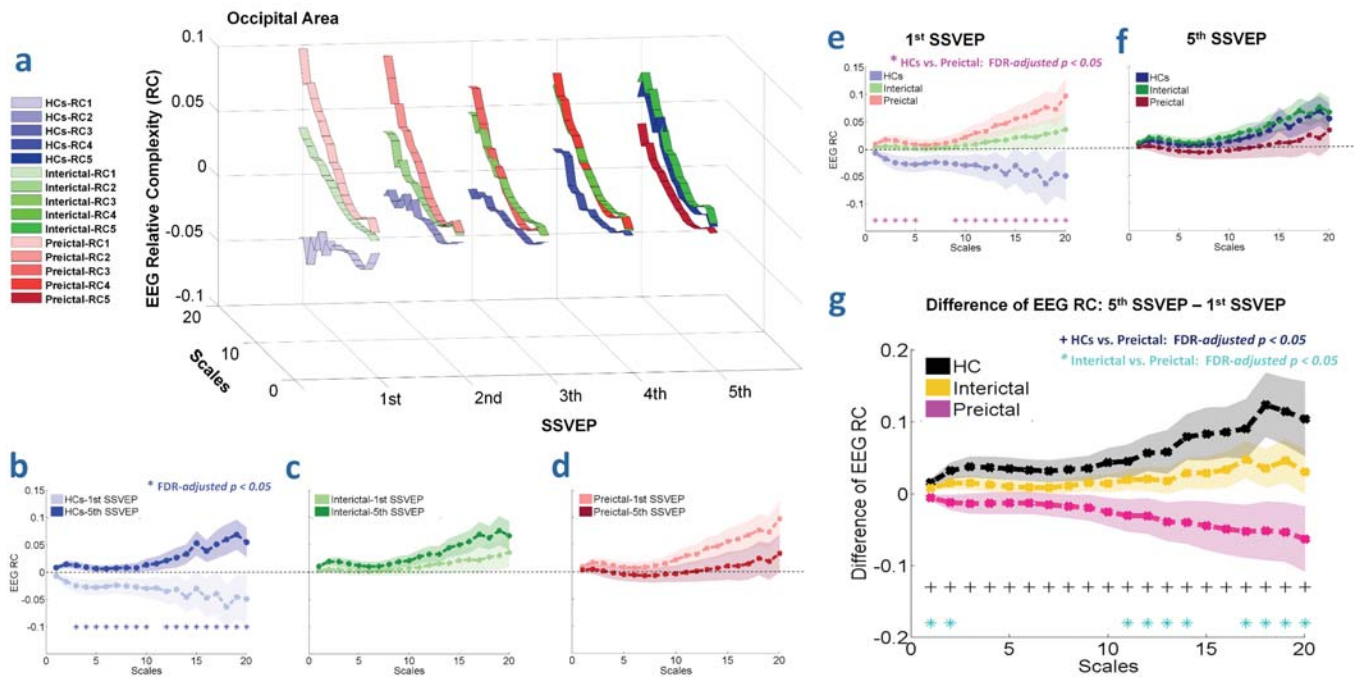
Figure 5 The trends in EEG entropy in the occipital area. (A) The changes in EEG entropy during the five (1st, 2nd, 3rd, 4th, and 5th marks) SSVEP stimuli over time scales ranging from 1 to 20 in HCs and migraine patients during the inter-ictal and pre-ictal phases. (B) Comparisons of EEG entropy between the 1st and 5th stimuli over time scales ranging from 1 to 20 in HCs. (C) Comparison of EEG entropy between the 1st and 5th stimuli over a time scale of 1 to 20 in patients of inter-ictal phase. (D) Comparisons of EEG entropy between the 1st and 5th stimuli over time scales ranging from 1 to 20 in patients during the pre-ictal phase. (E) Comparisons of EEG entropy among the inter-ictal phase, the pre-ictal phase and HCs over time scales ranging from 1 to 20 in the 1st stimuli session. (F) Comparisons of EEG entropy among the inter-ictal phase, the pre-ictal phase and HCs over time scales ranging from 1 to 20 in the 5th stimuli session. (G) Variance (5th stimuli EEG entropy minus 1st stimuli EEG entropy) comparisons among the inter-ictal phase, the pre-ictal phase and HCs. Of note, the traces represent the mean ± standard deviation (SD) of the EEG entropy of the inter-ictal phase, the pre-ictal phase or HCs. The asterisk or cross denotes a significant difference between different conditions (FDR-adjusted $p < 0.05$).

between different sessions (1st stimuli vs. 5th stimuli) in three groups of participants, i.e., HCs, inter-ictal patients, and pre-ictal patients. As shown in Fig. 5-B, EEG entropy showed a decreasing trend in entropy with an increasing time scale in the 1st stimulus session, which changed to an increasing trend in the 5th stimulus session in HCs. Furthermore, when comparing EEG entropy between the 1st and 5th stimulus sessions, the results showed an enhancement in occipital EEG entropy with increasing stimulus times for HCs. Table I shows the values of EEG entropy and p values in the 1st and 5th stimuli sessions. The EEG entropy increased from negative to positive values in the HCs group. The paired t-tests revealed that the EEG entropy in the 5th SSVEP session was significantly higher than that in the 1st SSVEP session in most time scales (FDR-adjusted $p < 0.05$).

For migraine patients in the inter-ictal phase (Fig. 5-C), EEG entropy had an ascending tendency with increasing time scales in the 1st and 5th stimulus sessions. Furthermore, EEG entropy had an increasing trend from the 1st to the 5th stimulus in all time scales, although the difference was not significant. However, for migraine patients in the pre-ictal phase (Fig. 5-D), although EEG entropy also retained an ascending tendency with increasing time scales in the 1st and 5th stimulus sessions, the entropy showed an opposite trend from the 1st to the 5th stimulus in all time scales. That is, EEG entropy showed a decreasing trend in the 5th stimuli session compared with the 1st stimuli session.

B. Inter-group comparisons of EEG entropy

In this section, we compared the differences in EEG entropy in the specific stimuli sessions among the three groups (HCs, inter-ictal patients and pre-ictal patients). During the 1st stimuli session (Fig. 5-E), the HCs group presented lower occipital EEG entropy relative to the migraine patients. The EEG entropy values of the HCs and the migraine patients were negative and positive, respectively, for all time scales. We did not find a

significant difference in EEG entropy between the HCs and the patients in the inter-ictal phase, but patients in the pre-ictal phase had significantly higher EEG entropy compared with the HCs (FDR-adjusted $p < 0.05$). In Table I, we present the values of EEG entropy and the p values for HCs and patients in the pre-ictal phase. In migraine patients, the pre-ictal phase was associated with a rising trend in EEG entropy at large time
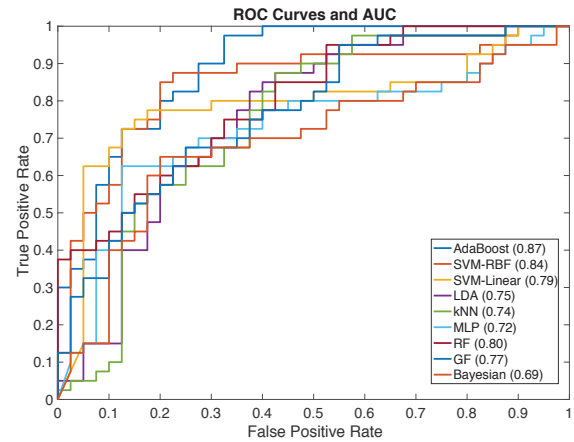

Figure 6 ROC curves and AUC are analyzed and plotted for each of the classifiers. The AUC is marked in the brackets of the legends (e.g., 0.87 AUC in AdaBoost).

scales compared with the inter-ictal phase, although there were no significant differences between these two phases.

By the 5th stimulus (Fig. 5-F), participants in all three groups exhibited positive entropy values in most time scales. Patients in the pre-ictal phase presented the lowest occipital EEG entropy relative to HCs and patients in the inter-ictal phase, although the EEG entropy did not significantly differ among the

Table I EEG entropy (mean (standard deviation)) in HCs and patients.

| τ | HC | | Pre-ictal | P^1 | P^2 |
	1st SSVEP	5th SSVEP	1st SSVEP		
1	-0.007 (0.005)	0.008 (0.003)	0.009 (0.003)	0.053	0.014
2	-0.018 (0.011)	0.014 (0.006)	0.017 (0.007)	0.059	0.012
3	-0.025 (0.013)	0.012 (0.006)	0.015 (0.009)	0.014	0.016
4	-0.027 (0.014)	0.009 (0.006)	0.011 (0.009)	0.025	0.028
5	-0.028 (0.014)	0.006 (0.006)	0.008 (0.010)	0.032	0.039
6	-0.026 (0.014)	0.006 (0.006)	0.007 (0.010)	0.037	0.057
7	-0.024 (0.014)	0.007 (0.006)	0.008 (0.010)	0.045	0.069
8	-0.026 (0.014)	0.008 (0.007)	0.011 (0.011)	0.043	0.057
9	-0.027 (0.016)	0.008 (0.009)	0.015 (0.012)	0.048	0.041
10	-0.030 (0.018)	0.013 (0.010)	0.021 (0.014)	0.039	0.029
11	-0.029 (0.020)	0.015 (0.013)	0.029 (0.016)	0.064	0.029
12	-0.035 (0.023)	0.021 (0.014)	0.032 (0.018)	0.030	0.024
13	-0.032 (0.026)	0.026 (0.018)	0.043 (0.019)	0.045	0.027
14	-0.046 (0.028)	0.033 (0.017)	0.046 (0.021)	0.018	0.012
15	-0.030 (0.033)	0.053 (0.022)	0.055 (0.023)	0.030	0.039
16	-0.047 (0.035)	0.038 (0.020)	0.060 (0.024)	0.019	0.014
17	-0.039 (0.039)	0.051 (0.025)	0.067 (0.026)	0.036	0.027
18	-0.064 (0.039)	0.060 (0.027)	0.076 (0.028)	0.008	0.005
19	-0.046 (0.039)	0.068 (0.027)	0.072 (0.029)	0.019	0.019
20	-0.049 (0.047)	0.054 (0.027)	0.097 (0.032)	0.049	0.012

τ represents the time scales.
[1] 1st SSVEP vs. 5th SSVEP for HCs.
[2] Pre-ictal patients vs. HCs in the 1st SSVEP session.

Table II Variance of EEG entropy (mean (standard deviation)) in HCs and patients

| τ | HC | Patients | | P^1 | P^2 |
		Inter-ictal	Pre-ictal		
1	0.015 (0.005)	0.007 (0.004)	-0.005 (0.004)	0.004	0.024
2	0.031 (0.011)	0.014 (0.008)	-0.012 (0.010)	0.006	0.035
3	0.037 (0.014)	0.014 (0.010)	-0.014 (0.014)	0.013	0.074
4	0.036 (0.015)	0.012 (0.010)	-0.013 (0.015)	0.024	0.129
5	0.034 (0.015)	0.009 (0.009)	-0.013 (0.015)	0.030	0.182
6	0.032 (0.015)	0.008 (0.009)	-0.013 (0.015)	0.036	0.208
7	0.031 (0.015)	0.008 (0.009)	-0.015 (0.016)	0.034	0.183
8	0.033 (0.016)	0.010 (0.010)	-0.017 (0.018)	0.028	0.125
9	0.034 (0.017)	0.014 (0.011)	-0.019 (0.019)	0.033	0.095
10	0.043 (0.020)	0.014 (0.012)	-0.025 (0.022)	0.016	0.081
11	0.044 (0.023)	0.019 (0.014)	-0.030 (0.024)	0.023	0.047
12	0.056 (0.025)	0.020 (0.010)	-0.031 (0.027)	0.015	0.049
13	0.057 (0.030)	0.017 (0.018)	-0.039 (0.029)	0.019	0.045
14	0.078 (0.032)	0.028 (0.021)	-0.040 (0.034)	0.008	0.048
15	0.082 (0.036)	0.028 (0.023)	-0.044 (0.035)	0.013	0.070
16	0.085 (0.035)	0.033 (0.025)	-0.049 (0.037)	0.008	0.050
17	0.090 (0.041)	0.047 (0.026)	-0.052 (0.040)	0.012	0.019
18	0.123 (0.044)	0.034 (0.027)	-0.051 (0.041)	0.005	0.044
19	0.113 (0.046)	0.045 (0.030)	-0.053 (0.041)	0.009	0.036
20	0.103 (0.052)	0.029 (0.031)	-0.063 (0.045)	0.018	0.049

τ represents the time scales.
[1] Pre-ictal patients vs. HCs.
[2] Pre-ictal vs. inter-ictal phases.

three groups. Of note, EEG entropy of HCs is similar to that of patients of inter-ictal phase in most time scales.

Additionally, we defined the transitional variance of EEG entropy as the difference in EEG entropy between the 1st and 5th stimulus sessions. As shown in Fig. 5-G, patients in the pre-ictal phase exhibited a significantly lower transitional variance of EEG entropy than that in the inter-ictal phase for most large time scales. Similarly, patients in the pre-ictal phase had a significantly lower transitional variance of EEG entropy than the HCs for all time scales (FDR-adjusted $p < 0.05$). Table II shows the transitional variance of EEG entropy and the p values for HCs and patients in the inter-ictal and pre-ictal phases. Considering the transitional variance of EEG entropy in individuals from the inter-ictal to pre-ictal phase, for the time scale of 20, entropy decreased in 28 of the 40 patients (70%) when they entered the pre-ictal phase from the inter-ictal phase. In contrast, an increment in entropy was observed in 12 patients. Additionally, 8 patients with two pre-ictal examinations were selected for the test-retest reliability. According to the quoted guidelines for interpretation of inter-rater agreement measures, our results showed a good reliability with an intra-class correlation coefficient $r1$ of 0.70 ($p = 0.04$).

Table III Performances of the classification models

		Accuracy	Recall	Precision	F-measure
STATISTICS	LDA	0.72 (0.05)*	0.69 (0.05)*	0.72 (0.05)*	0.71 (0.05)*
	kNN	0.70 (0.06)*	0.68 (0.06)*	0.70 (0.06)*	0.69 (0.06)*
	MLP	0.69 (0.07)*	0.63 (0.05)*	0.67 (0.06)*	0.66 (0.07)*
	Bayesian	0.67 (0.06)*	0.63 (0.06)*	0.67 (0.06)*	0.67 (0.06)*
	SVM Linear	0.74 (0.05)*	0.71 (0.05)*	0.75 (0.06)*	0.73 (0.05)*
	SVM RBF	0.79 (0.05)	0.78 (0.06)	0.77 (0.05)	0.76 (0.05)
	GP	0.73 (0.05)*	0.70 (0.05)*	0.72 (0.05)*	0.72 (0.05)*
	RF	0.75 (0.06)*	0.73 (0.05)*	0.75 (0.06)*	0.74 (0.05)*
	AdaBoost	**0.81 (0.06)**	**0.80 (0.06)**	**0.79 (0.06)**	**0.78 (0.06)**
PCA	LDA	0.71 (0.04)*	0.70 (0.05)*	0.71 (0.05)*	0.71 (0.04)*
	kNN	0.69 (0.05)*	0.69 (0.07)*	0.69 (0.07)*	0.70 (0.07)*
	MLP	0.65 (0.06)*	0.64 (0.07)*	0.63 (0.06)*	0.64 (0.06)*
	Bayesian	0.63 (0.07)*	0.61 (0.07)*	0.64 (0.06)*	0.64 (0.07)*
	SVM Linear	0.72 (0.05)*	0.70 (0.05)*	0.73 (0.06)*	0.71 (0.05)*
	SVM RBF	0.78 (0.06)	0.77 (0.06)	0.76 (0.06)	0.75 (0.06)
	GP	0.72 (0.06)*	0.69 (0.06)*	0.70 (0.07)*	0.70 (0.06)*
	RF	0.73 (0.06)*	0.72 (0.06)*	0.74 (0.06)*	0.73 (0.06)*
	AdaBoost	**0.80 (0.05)**	**0.78 (0.06)**	**0.78 (0.05)**	**0.77 (0.06)**

Parameters: LDA: default settings; KNN: k=3; MLP: structure with one hidden layer and number of units=5; Bayesian: default settings; SVM with linear kernel: c=0; SVM with RBF kernel: c=15, g=10; GP: exponential covariance [-0.73 -0.12] and Gaussian likelihood -2.58; RF: the number of decision trees n=20; AdaBoost: the number of weak classifiers n=8 (LDA, kNN, MLP, Bayesian, Linear-SVM, RBF-SVM, GP, and RF).

*: After utilizing one-way ANOVA, Tukey's post hoc test was performed for pairwise comparisons: $p < 0.05$ (AdaBoost vs. Other Classifiers).

Table IV. Root Mean Square Deviation (RMSE) Evaluation
(Mean + Standard Deviation)

Scale	RMSE of Entropy Values Exponential	RMSE of Entropy Values Gaussian	Scale	RMSE of Entropy Values Exponential	RMSE of Entropy Values Gaussian
1	0.015±0.004	0.016±0.004	11	0.015±0.005	0.016±0.005
2	0.018±0.005	0.018±0.005	12	0.016±0.005	0.016±0.005
3	0.015±0.004	0.016±0.004	13	0.016±0.006	0.016±0.006
4	0.018±0.004	0.019±0.004	14	0.016±0.005	0.016±0.005
5	0.015±0.004	0.015±0.004	15	0.015±0.005	0.015±0.005
6	0.016±0.007	0.016±0.007	16	0.017±0.006	0.017±0.006
7	0.014±0.005	0.014±0.005	17	0.016±0.005	0.016±0.005
8	0.015±0.005	0.015±0.005	18	0.016±0.005	0.017±0.005
9	0.014±0.006	0.014±0.006	19	0.017±0.005	0.017±0.006
10	0.013±0.004	0.014±0.004	20	0.017±0.006	0.018±0.006

C. Performance of classification models

To capture useful input variables, we selected features (transitional variance of EEG entropy) for specific time scales with significant changes. Thus, the transitional variance of EEG entropy with time scales of $\tau = 1, 2, 11, 12, 13, 14, 17, 18, 19, 20$ were chosen as the input features of the classification models. Furthermore, we added the principal component analysis (PCA) as a feature selection scheme and the leave-one-out cross-validation procedure to identify the best features and evaluate the ability of generalized classification. Here, we choose the 10 eigen values in PCA to be consistent with the statistical analysis.

As shown in Table III, the performances of six binary classification models (LDA, kNN, MLP, Bayesian, Linear/RBF-SVM, GP, RF, and AdaBoost) for classifying migraine phases (inter-ictal vs. pre-ictal) were evaluated. Furthermore, we selected four matrices (accuracy, recall, precision, and F-measure) to assess classification performance. Considering the feature selection with statistical scheme, in terms of accuracy and recall, the AdaBoost ensemble learning classifier (81±6% accuracy and 80±6% recall) significantly outperformed other classifiers ($p < 0.05$) except for the RBF-SVM. In terms of precision and F-measure, the AdaBoost ensemble learning classifier (79±6% precision and 78±5% F-measure) had comparable results that were significantly better than the results obtained with the LDA, kNN, MLP, Bayesian, Linear-SVM, GP and RF classifiers ($p < 0.05$). Similarly, considering the feature selection with PCA scheme, the AdaBoost ensemble learning classifier showed the best performance with 80±5% accuracy, 78±6% recall, 78±5% precision, and 77±6% F-measure.

Additionally, the ROC curves and AUC were employed to evaluate the performance of various classifiers with feature
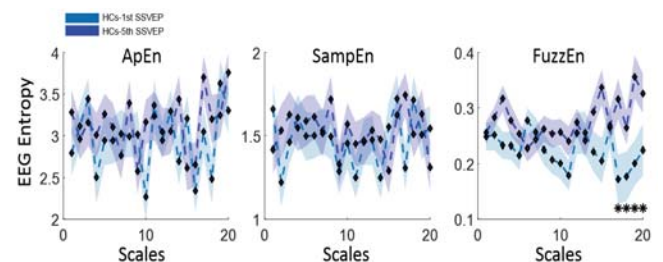

Figure 7 Comparisons of EEG entropy between the 1st and 5th stimuli over 1 to 20 time scales in HCs as measured by ApEn, SampEn and FuzzEn.

selection of statistical analysis. Figure 6 showed the ROC curves and AUC of the classifiers to compare the classification performance among various classifiers. The horizontal axis and vertical axis represented the false positive rate and true positive rate of classifiers, respectively. Each curve line indicated ROC of classifiers, and the areas under curve lines presented the AUC of classifiers. Of note, AdaBoost had the highest AUC of 0.87, and SVMRBF had the second-highest AUC of 0.84 in estimating the migraine phases.

D. Performance of competing entropy algorithms

In this section, we used the same EEG processing steps but estimated entropy using the competing entropy algorithms (ApEn, SampEn, and FuzzEn), of which, we used the same parameters ($m = 2, r = 0.15$ and $\tau = 20$) to be consistent with our proposed method. In HCs (Fig. 7), we could not distinguish EEG entropy between the 1st and 5th stimuli sessions using ApEn or SampEn without a fuzzy structure. However, the FuzzEn algorithm could only distinguish the EEG entropy of these two sessions at larger time scales ($\tau = 17, 18, 19, 20$), suggesting that the fuzzy structure (including FuzzEn and Inherent FuzzEn) is better than the non-fuzzy structure (including ApEn and SampEn) in estimating EEG entropy in SSVEPs. Furthermore, there was no significant difference in EEG entropy between each session in migraine patients, indicating that the relative inherent fuzzy entropy algorithm is superior to the other competing models, including ApEn, SampEn and FuzzEn, for use in SSVEP-based migraine studies.

Additionally, the fuzzy membership function that possesses the following desired properties can be chosen: a. being continuous so that the similarity does not change abruptly; b. being convex so that self-similarity is the maximum. For example, we chose a resting-state EEG signal with 3 closed-eye epochs (Fpz channel), featuring by similar entropy values and minimized deviations from each other. Followed-up our previous study, we used healthy subjects and applied root mean square deviation (RMSD) $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i^{(\tau)} - \widehat{y^{(\tau)}})^2}$ to evaluate which fuzzy membership function achieved better performance (Exponential vs. Gaussian). Here, $y_i^{(\tau)}$ indicates the entropy value in each epoch with certain scale τ, and $\hat{y}$ is the mean value of entropy among 3 closed-eyes epochs, and $n = 3$.

Table IV shows the RMSE of entropy values using exponential and Gaussian membership functions. In terms of EEG signals, it showed that entropy values in some scales measured with exponential membership function achieved smaller RMSE, compared to that measured with Gaussian membership function. The results indicates that exponential membership function has better performance than Gaussian membership function in terms of EEG signals; however, generally speaking, choosing different fuzzy membership functions does not influence significantly on entropy values.

V. DISCUSSION

A. Complexity characteristics

Systemic analyses indicate that entropy dynamics reflect the nonlinear complex characteristics of the brain that allow it to adapt to constantly changing stimulus situations rather than the linear characteristics of the brain [49]. The associated entropy models, which can be used to quantify brain complexity to express the robustness of brain systems, are crucial for quantifying the critical characteristics of nonlinear neuro-dynamics [50]. Additionally, as biological systems functioning on different time scales may exhibit different behaviors, so bio-signals are often examined at multiple time scales. Our results showed that EEG complexity was enhanced in the occipital area after repetitive visual stimulations at different time scales, which was in line with the finding of a previous study showing that EEG entropy exhibited an increasing trend in response to long-term audio-visual stimulation [51].

On the other hand, healthy individuals learn to stop responding to stimuli that are no longer biologically relevant [52, 53]. For example, healthy humans can habituate to repeated visual stimuli when they learn these have no consequences. This habituation performance is a form of adaptive behavior that reflects the robustness of brain systems. Our finding that EEG complexity was significantly enhanced during repetitive SSVEPs in HCs is considered another form of habitation indicating the robustness of brain systems. Traditionally researchers who study habituation have focused on a time/frequency dimension of the behavior (i.e. response potential or magnitude). Our high throughput entropy analyses of habituation for complexity estimation have changed this view. Our study offers a new way to think about the role of inherent fuzzy entropy in the context of adaptive behavioral strategies, i.e., habituation.

B. EEG complexity in migraine patients

Prior studies have noted that healthy brain system present robustness with high complexity but diseased brain systems show vulnerable characteristics with low complexity [20, 29]. Healthy subjects exhibit normal habituation to repetitive visual stimulation [30] but dishabituation has been observed in migraine patients [6-11]. Our study discovered changes in EEG complexity in migraine patients by using a multi-scale relative inherent fuzzy entropy algorithm in a repetitive SSVEP environment. We found that occipital complexity was slightly, although not significant, enhanced with increasing stimulus times for migraine patients in the inter-ictal phase, but a reversed trend was observed for migraine patients in the pre-ictal phase. Thus, our findings are consistent with the prior studies and may suggest that migraine patients in the pre-ictal phase are more vulnerable to adapt a repetitive visual simulation environment than those in the inter-ictal phase.

Regarding the values of relative inherent fuzzy entropy in the 1st SSVEP session, a key result is that migraine patients and HCs have positive and negative entropy values, respectively. This finding suggests that migraine patients and HCs have higher and lower EEG complexities, respectively, relative to baseline (resting-state) when responding to the 1st visual stimulation.

Additionally, it would be, nevertheless, of great interest to evaluate the EEG complexity by using different stimulation frequencies, either lower or higher ones. Since previous SSVEP studies all demonstrated a hyper-responsive pattern in patients with migraine throughout a low- to median-to-high frequency range, we speculate the results would be similar. Further studies are required to confirm this point.

C. EEG-based headband BCI migraine system

Because of its broad availability and cost-effectiveness, EEG is widely used as a non-invasive means to assess dynamic changes in brain electrical activity. The rapid development of dry sensors and wearable devices [36, 54, 55] has led to a reduction in the preparatory work required for long-term monitoring. Moreover, the headband design with occipital electrodes is convenient for long-term monitoring and daily use [36, 54, 55]. Therefore, it is possible to implement EEG-based models in laboratories and real-world settings.

Considering the transitional variance in EEG entropy, our findings showed that migraine patients in the pre-ictal phase presented significantly lower occipital complexity than migraine patients in the inter-ictal phase or HCs. Due to the distinctive specifications of each classifier, we used 9 EEG-based classifiers with the input of the transitional variance in EEG entropy, to test and compare the performance on which of two migraine phases a new EEG signal belongs. Of them, the SVM constructs a hyperplane in a high-dimensional space, which has the largest distance to the nearest training-data point of any class. The GF is a probabilistic classification model specified by a mean function and a covariance function. The RF is basically an ensemble of decision trees, and each tree classifies the dataset using a subset of variables. Especially, the AdaBoost, a machine learning meta-algorithm, can be used in conjunction with many other types of classifiers to improve the performance. In the study, the output of the other classifiers including LDA, kNN, MLP, Bayesian, Linear-SVM, RBF-SVM, GP, and RF, called weak classifiers, is combined into a weighted sum that represents the final output of the boosted classifier. This outstanding outcome contributed to discriminating between inter-ictal and pre-ictal migraine phases with 81±6% accuracy, 80±6% recall, 79±6% precision and 78±6% F-measure, as well as the AUC of 0.87.

In a 2-class decoding problem with a small set (n=80), the classification performance significantly exceeds chance if the accuracy reaches 70% ($p < 10^{-4}$) [56]. In our study, the accuracy of classifiers is larger than 70%, which indicated that the performance of our classifiers exceeded the chance level. Thus, the wearable EEG solution and the characteristics of multi-scale relative inherent fuzzy entropy can be easily assessed in repetitive visual stimulation paradigms. However, the transitional variance of EEG entropy indeed affects the performance of classifiers to obtain a good generalization of validation. What is more, the computational time of the multi-scale inherent fuzzy entropy algorithm prevents online applications or even the processing of long data sets, due to the computational time of the fuzzy entropy value. Our next step will conduct a follow-up study to record daily EEG signals in out-patient individuals to evaluate the transitional variance between the two migraine phases, and calculate entropy values with less computational time. In the future, we believe that occipital complexity features can be evaluated to develop a brain-computer interface (BCI) system that can be used to recognize migraine phases before a migraine attack.

D. Advantages of inherent fuzzy entropy

Despite recent progress that has been made in visual stimulus research, extracting corticocerebral complexity using EEG-based multi-scale entropy approaches that can be used to determine the robustness of brain systems remains challenging.

In previous studies, researchers have generally ignored the superimposed trends in EEG signals, which leads to poor performance of entropy algorithms in realistic EEG applications. Thus, we developed a multi-scale inherent fuzzy entropy algorithm [28] that has been applied in a previous resting-state migraine study [29]. In this SSVEP-based study, we modified the original version of the algorithm to a multi-scale relative inherent fuzzy entropy algorithm aiming to eliminate individual differences.

In addition to testing the Inherent FuzzyEn algorithm, we employed the same data processing steps using other entropy algorithms, one with a fuzzy structure (FuzzEn) and one with a non-fuzzy structure (SampEn and ApEn). Considering healthy subjects who received repetitive visual stimulations, our findings showed that the entropy algorithms with fuzzy structures (Inherent FuzzyEn and FuzzEn) exhibited good performance compared with algorithms with non-fuzzy structures (SampEn and ApEn). Furthermore, the performance of the Inherent FuzzyEn algorithm was superior to the performances of the FuzzyEn, SampEn and ApEn models. In summary, the Inherent FuzzyEn algorithm was more effective in evaluating EEG signals from healthy subjects and migraine patients not only during resting-state but also during the repetitive SSVEP condition.

E. Future perspectives

In this study, we have demonstrated the feasibility of the SSVEP-based complexity using a wearable headband EEG in support of detection of migraine attacks. The use of dry electrodes allows for easy and rapid monitoring on a daily basis and the advances in EEG recording and analysis ensure a promising future in support of individual solutions. The following research aims to improve the inherent fuzzy entropy model for online applications with EEG signals. We have ongoing advances in practical approaches of brain signal recording, and sophisticated designs of extracting knowledge from neuro-information and home healthcare solutions are envisioned to guide a wide range of real-life applications in the near future. Additionally, we hope to extend the outlook for future research on representing brain dynamics under uncertainties, such as exploring stochastic maximum principles and Markov regime-switching jump-diffusion model with minimum delay [57], and loop-based conic multivariate adaptive model for advanced construction of complexity networks [58].

VI. CONCLUSION

This study extracted SSVEP-based multi-scale relative inherent fuzzy entropy from migraine patients and HCs using a wearable headband EEG device. Our results highlight the feasibility of using a novel entropy measurement to compare EEG complexity in repetitive visual stimulation. Occipital EEG entropy showed an enhanced trend in patients in the inter-ictal phase; whereas, a reverse trend in patients in the pre-ictal phase. Additionally, patients in the pre-ictal phase exhibited a significantly lower transitional variation in EEG entropy than patients in the inter-ictal phase. Of note, we also noted that inherent fuzzy entropy was superior to other competing entropy models for conducting SSVEP experiments. In summary, inherent fuzzy entropy can be used in novel applications of

visual stimulus environments for migraine studies, which may potentially be used in the future to provide a pre-ictal alert to migraine patients.

REFERENCES

[1] Headache Classification Committee of the International Headache Society (IHS), "The international classification of headache disorders, (beta version)," *Cephalalgia,* vol. 33, no. 9, pp. 629-808, 2013.

[2] M. Bjørk, L. Stovner, B. Nilsen, M. Stjern, K. Hagen, and T. Sand, "The occipital alpha rhythm related to the "migraine cycle" and headache burden: a blinded, controlled longitudinal study," *Clin. Neurophysiol.,* vol. 120, no. 3, pp. 464-471, 2009.

[3] A. M. Norcia, L. G. Appelbaum, J. M. Ales, B. R. Cottereau, and B. Rossion, "The steady-state visual evoked potential in vision research: a review," *J. Vis.,* vol. 15, no. 6, pp. 4-4, 2015.

[4] P. Sehatpour *et al.*, "A human intracranial study of long-range oscillatory coherence across a frontal–occipital–hippocampal brain network during visual object processing," *Proc. Natl. Acad. Sci.,* vol. 105, no. 11, pp. 4399-4404, 2008.

[5] V. Romei, V. Brodbeck, C. Michel, A. Amedi, A. Pascual-Leone, and G. Thut, "Spontaneous fluctuations in posterior α-band EEG activity reflect variability in excitability of human visual areas," *Cereb. Cortex,* vol. 18, no. 9, pp. 2010-2018, 2007.

[6] G. Coppola, F. Pierelli, and J. Schoenen, "Habituation and migraine," *Neurobiol. Learn. Mem.,* vol. 92, pp. 249-259, 2009.

[7] M. J. Mickleborough, C. M. Toma, Jeremy, A. S. Chan, J. H. Truong and T. C. Handy, "Interictal Neurocognitive Processing of Visual Stimuli in Migraine: Evidence from Event-Related Potentials," *PLoS One,* vol. 8, no. 11, 2013.

[8] S-J. Wang. W-T. Chen, J-L. Fuh, C-P. Lin, Y-C. Ko and Y-Y. Lin, "Peri-ictal normalization of visual cortex excitability in migraine: an MEG study." *Cephalalgia,* vol. 29, pp. 1201-1211, 2009.

[9] A. Judit, P. S. Sandor, and J. Schoenen, "Habituation of visual and intensity dependence of auditory evoked cortical potentials tends to normalize just before and during the migraine attack.," *Cephalalgia* vol. 20, pp. 714-19, 2000.

[10] L. C. Beese, D. Putzer, N. Osada, N. Evers and M. Marziniak, "Contact heat evoked potentials and habituation measured interictally in migraineurs," *J. Headache Pain,* vol. 16, no. 1, 2015.

[11] G. Cosentino, *et al.*, "Cyclical changes of cortical excitability and metaplasticity in migraine: evidence from a repetitive transcranial magnetic stimulation study.," *Pain,* vol. 155, no. 6, p. 8, 2014.

[12] S. Genco, M. Tommaso, A. Prudenzano, M. Savarese, and F. Puca, "EEG features in juvenile migraine: topographic analysis of spontaneous and visual evoked brain electrical activity: a comparison with adult migraine," *Cephalalgia,* vol. 14, no. 1, pp. 41-46, 1994.

[13] F. Puca, M. Tommaso, P. Tota, and V. Sciruicchio, "Photic driving in migraine: correlations with clinical features," *Cephalalgia,* vol. 16, no. 4, pp. 246-250, 1996.

[14] M. Tommaso *et al.*, "Steady-state visual evoked potentials in the low frequency range in migraine: a study of habituation and variability phenomena," *Int. J. Psychophysiol.,* vol. 49, no. 2, pp. 165-174, 2003.

[15] R. S. Fisher, G. Harding, G. Erba, G. L. Barkley, and A. Wilkins, "Photic-and pattern-induced seizures: a review for the Epilepsy Foundation of America Working Group," *Epilepsia,* vol. 46, no. 9, pp. 1426-1441, 2005.

[16] H. Wang *et al.*, "Discriminative feature extraction via multivariate linear regression for SSVEP-based BCI," *IEEE Trans. Neural. Syst. Rehabil. Eng.,* vol. 24, no. 5, pp. 532-541, 2016.

[17] Y. 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. 04, pp. 1450013, 2014.

[18] Y. Jiao, Y. Zhang, Y. Wang, B. Wang, J. Jin, and X. Wang, "A novel multilayer correlation maximization model for improving CCA-based frequency recognition in SSVEP brain–computer interface," *Int. J. Neural Syst.,* vol. 28, no. 04, pp. 1750039, 2018.

[19] Y. Zhang, G. Zhou, J. Jin, M. Wang, X. Wang, and A. Cichocki, "L1-regularized multiway canonical correlation analysis for SSVEP-based BCI," *IEEE Trans. Neural. Syst. Rehabil. Eng.,* vol. 21, no. 6, pp. 887-896, 2013.

[20] T. Takahashi *et al.*, "Antipsychotics reverse abnormal EEG complexity in drug-naive schizophrenia: a multiscale entropy analysis," *Neuroimage,* vol. 51, no. 1, pp. 173-182, 2010.

[21] J. Gao, J. Hu, F. Liu, and Y. Cao, "Multiscale entropy analysis of biological signals: a fundamental bi-scaling law," *Front. Comput. Neurosci.,* vol. 9, 2015.

[22] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proc. Natl. Acad. Sci.,* vol. 88, no. 6, pp. 2297-2301, 1991.

[23] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Am. J. Physiol. Heart Circ. Physiol.,* vol. 278, no. 6, pp. 2039-2049, 2000.

[24] I. J. Rudas and M. O. Kaynak, "Entropy-based operations on fuzzy sets," *IEEE Trans. Fuzzy Syst.,* vol. 6, no. 1, pp. 33-40, 1998.

[25] W. Chen, Z. Wang, H. Xie, and W. Yu, "Characterization of surface EMG signal based on fuzzy entropy," *IEEE Trans. Neural. Syst. Rehabil. Eng.,* vol. 15, no. 2, pp. 266-272, 2007.

[26] P. Li and B. Liu, "Entropy of credibility distributions for fuzzy variables," *IEEE Trans. Fuzzy Syst.,* vol. 16, no. 1, pp. 123-129, 2008.

[27] H-B. Xie, B. Sivakumar, T. W. Boonstra, and K. Mengersen, "Fuzzy Entropy and Its Application for Enhanced Subspace Filtering," *IEEE Trans. Fuzzy Syst.,* 2017.

[28] Z. Cao and C-T. Lin, "Inherent Fuzzy Entropy for the Improvement of EEG Complexity Evaluation," *IEEE Trans. Fuzzy Syst.,* 2018.

[29] Z. Cao, K-L. Lai, C-T. Lin, C-H. Chuang, C-C. Chou, and S-J. Wang, "Exploring resting-state EEG complexity before migraine attacks," *Cephalalgia,* vol. 38, no. 7, pp. 1296-1306, 2018.

[30] M. E. Bouton, *Learning and behavior: A contemporary synthesis*. Sinauer Associates, 2007.

[31] R. Mahajan and B. I. Morshed, "Unsupervised eye blink artifact denoising of EEG data with modified multiscale sample entropy, kurtosis, and Wavelet-ICA," *IEEE J. Biomed. Health Inform.,* vol. 19, no. 1, pp. 158-165, 2015.

[32] H-B. Xie, W-X. He, and H. Liu, "Measuring time series regularity using nonlinear similarity-based sample entropy," *Phys. Lett. A,* vol. 372, no. 48, pp. 7140-7146, 2008.

[33] P. Grassberger and I. Procaccia, "Estimation of the Kolmogorov entropy from a chaotic signal," *Phys. Rev. A,* vol. 28, no. 4, pp. 2591, 1983.

[34] M. Gerwig, L. Niehaus, O. Kastrup, P. Stude, and H. C. Diener, "Visual cortex excitability in migraine evaluated by single and paired magnetic stimuli," *Headache: J. Head and Face Pain,* vol. 45, no. 10, pp. 1394-1399, 2005.

[35] M. Bjørk, L. Stovner, K. Hagen, and T. Sand, "What initiates a migraine attack? Conclusions from four longitudinal studies of quantitative EEG and steady-state visual-evoked potentials in migraineurs," *Acta Neurol. Scand.,* vol. 124, no. 191, pp. 56-63, 2011.

[36] C-T. Lin, L-D. Liao, Y-H. Liu, I-J. Wang, B-S. Lin, and J-Y. Chang, "Novel dry polymer foam electrodes for long-term EEG measurement," *IEEE Trans. Biomed. Eng.,* vol. 58, no. 5, pp. 1200-1207, 2011.

[37] R. J. Barry, A. R. Clarke, S. J. Johnstone, C. A. Magee, and J. A. Rushby, "EEG differences between eyes-closed and eyes-open resting conditions," *Clin. Neurophysiol.,* vol. 118, no. 12, pp. 2765-2773, 2007.

[38] Y-P. Lin, Y. Wang, and T-P. Jung, "A Mobile SSVEP-based Brain-computer Interface for Freely Moving Humans: The Robustness of Canonical Correlation Analysis to Motion Artifacts," Annual International Conference of the IEEE EMBS, Osaka, Japan, 2013, pp. 1350-1353.

[39] 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, 2018.

[40] 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, pp. 046, 2015.

[41] C-T. Lin, C-S. Huang, W-Y. Yang, A. K. Singh, C-H. Chuang, and Y-K. Wang, "Real-Time EEG Signal Enhancement Using Canonical Correlation Analysis and Gaussian Mixture Clustering," *J. Healthc. Eng.,* vol. 2018, pp. 5081258, 2018.

[42] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 22, no. 1, pp. 4-37, 2000.

[43] C. E. Rasmussen and H. Nickisch, "Gaussian processes for machine learning (GPML) toolbox," *J. Mach. Learn. Res.,* vol. 11, no. 11, pp. 3011-3015, 2010.

[44] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R News,* vol. 2, no. 3, pp. 18-22, 2002.

[45] G. Rätsch, T. Onoda, and K-R. Müller, "Soft margins for AdaBoost," *Mach. Learn.,* vol. 42, no. 3, pp. 287-320, 2001.

[46] F. Heijden, R. Duin, D. d. Ridder, and D. M. J. Tax, *Classification, Parameter Estimation and State Estimation: An Engineering Approach Using MATLAB.* Hoboken, New Jersey: John Wiley & Sons, 2004.

[47] C-C. Chang and C-J. Lin, "LIBSVM: A Library for support vector machines," *ACM Trans. Intell. Syst. Technol.,* vol. 2, no. 3, pp. 1-27, 2011.

[48] C. Goutte and E. Gaussier, "A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation," in *Advances in Information Retrieval: 27th European Conference on IR Research,* Springer Berlin Heidelberg, 2005, pp. 345-359.

[49] M. Gazzaniga, K. Doron, and C. Funk, "Looking toward the future: Perspectives on examining the architecture and function of the human brain as a complex system," *Cog. Neurosci. IV,* pp. 1245-1252, 2010.

[50] J. Gao, J. Hu, and W-W. Tung, "Complexity measures of brain wave dynamics," *Cogn. Neurodyn.,* vol. 5, no. 2, pp. 171-182, 2011.

[51] M. Teplan, A. Krakovska, and S. Štolc, "EEG responses to long-term audio–visual stimulation," *Int. J. Psychophysiol.,* vol. 59, no. 2, pp. 81-90, 2006.

[52] P. M. Groves and R. F. Thompson, "Habituation: a dual-process theory," *Psychological review,* vol. 77, no. 5, pp. 419, 1970.

[53] R. F. Thompson and W. A. Spencer, "Habituation: a model phenomenon for the study of neuronal substrates of behavior," *Psychol. Rev.,* vol. 73, no. 1, pp. 16, 1966.

[54] C-T. Lin *et al.*, "Noninvasive neural prostheses using mobile and wireless EEG," *Proc. IEEE,* vol. 96, no. 7, pp. 1167-1183, 2008.

[55] L-D. Liao *et al.*, "Biosensor technologies for augmented brain–computer interfaces in the next decades," *Proc. IEEE,* vol. 100, no. Special Centennial Issue, pp. 1553-1566, 2012.

[56] E. Combrisson and K. Jerbi, "Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy," *J. Neurosci. Methods,* vol. 250, pp. 126-136, 2015.

[57] E. Savku and G. W. Weber, "A stochastic maximum principle for a markov regime-switching jump-diffusion model with delay and an application to finance," *J. Optimiz. Theory App.,* vol. 179, no. 2, pp. 696-721, 2018.

[58] E. Ayyıldız, V. Purutçuoğlu, and G. W. Weber, "Loop-based conic multivariate adaptive regression splines is a novel method for advanced construction of complex biological networks," *European J. Oper. Res.,* vol. 270, no. 3, pp. 852-861, 2018.

Zehong Cao is a Lecturer (a.k.a. Assistant Professor) with Discipline of Information and Communication Technology (ICT), School of Technology, Environments and Design, College of Sciences and Engineering, University of Tasmania, Hobart, Australia, and a Honorary Research Associate with Faculty of Engineering and IT, University of Technology Sydney, Australia. He received the dual PhD degree in Information Technology from University of Technology Sydney, and Electrical and Control Engineering from National Chiao Tung University. He received the MS and BS from The Chinese University of Hong Kong and Northeastern University, respectively. He serves as the Associate Editor of IEEE Access and the Guest Editor of Swarm and Evolutionary Computation, Neurocomputing, and International Journal of Distributed Sensor Networks. He was awarded UTS CAI Best Paper (2017), UTS FEIT Publication Award (2017), UTS President Scholarship (2015), and NCTU & Songshanhu Scholarship (2014). His research interests cover signal processing, brain-computer interface, data mining, bioinformatics, fuzzy systems, neural networks, pattern recognition, and clinical applications.

Chin-Teng Lin received a BS degree from National Chiao Tung University (NCTU), Taiwan, in 1986 and Master's and PhD degrees in electrical engineering from Purdue University, USA, in 1989 and 1992, respectively. He is currently a Distinguished Professor of Faculty of Engineering and Information Technology, University of Technology Sydney, and a Chair Professor of Electrical and Computer Engineering, NCTU, International Faculty of University of California at San-Diego (UCSD), and he holds an Honorary Professorship of University of Nottingham.

Dr. Lin was elevated to an IEEE Fellow for his contributions to biologically inspired information systems in 2005 and was elevated to an International Fuzzy Systems Association (IFSA) Fellow in 2012. He was elected as the Editor-in-chief of IEEE Transactions on Fuzzy Systems in 2011-2016. He also served on the Board of Governors at IEEE Circuits and Systems (CAS) Society in 2005-2008, IEEE Systems, Man, Cybernetics (SMC) Society in 2003-2005, IEEE Computational Intelligence Society (CIS) in 2008-2010, and Chair of IEEE Taipei Section in 2009-2010. Dr. Lin was a Distinguished Lecturer of IEEE CAS Society from 2003 to 2005 and CIS Society from 2015-2017. He served as the Deputy Editor-in-Chief of IEEE Transactions on Circuits and Systems-II in 2006-2008. Dr. Lin was the Program Chair of IEEE International Conference on Systems, Man, and Cybernetics in 2005 and General Chair of 2011 IEEE International Conference on Fuzzy Systems. Dr. Lin is the coauthor of Neural Fuzzy Systems (Prentice-Hall) and the author of Neural Fuzzy Control Systems with Structure and Parameter Learning (World Scientific). He has published more than 200 journal papers (Total Citation: 17,687, H-index: 61, i10-index: 217) in the areas of neural networks, fuzzy systems, multimedia hardware/software, and cognitive neuro-engineering, including approximately 120 IEEE journal papers.

Kuan-Lin Lai received an MD degree from the School of Medicine, National Yang-Ming University (NYMU), Taipei, Taiwan, in 2002. Currently, he is an attending physician at the Neurological Institute, Taipei Veterans General Hospital and a conjunct instructor at National Yang-Ming University School of Medicine, Taipei, Taiwan. He was awarded the second prize of best poster award at the 19th Taiwan Neurological Society Congress (2014) and the Young Investigator award at the 29th International Congress of Clinical Neurophysiology (2010). He has published over 20 journal and conference papers.

Li-Wei Ko received the BS degree in Mathematics from National Chung Cheng University, Chiayi, Taiwan, the MS degree in educational measurement and statistics from the National Taichung University of Education, Taichung City, Taiwan, and the PhD degree in Electrical Engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2001, 2004, and 2007, respectively. Currently, he is an Associate Professor in the Department of Biological Science and Technology, and the Brain Research Center at National Chiao Tung University, Taiwan. His research interests include brain–computer interface, computational neuroscience, neural rehabilitation, neural fuzzy systems and machine learning.

Jung-Tai King received the BS degree in Psychology from the National Cheng-Chi University in 1998, the MS degree in criminology from the National Chung-Cheng University in 2001, and the PhD degree in Neuroscience from National Yang-Ming University, Taiwan in 2010. He is currently an Assistant Research Fellow in the Brain Research Center, National Chiao Tung University, Taiwan. His research interests include psychophysiology, cognitive and social neuroscience, and neuro-marketing.

Kwong-Kum Liao received an MD degree from the School of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan, in 1982. Currently, he is an attending physician at the Neurological Institute, Taipei Veterans General Hospital. His major interest includes electrophysiology, human reflex responses, sensorimotor integration in motor control, convergence and divergence of sensory inputs.

Jong-Ling Fuh completed her medical education at the National Yang-Ming University in Taipei, and went on to do her Behavior Neurology Fellowship at UCLA, USA, in 1996. She is currently a Professor of Neurology at the National Yang-Ming University School of Medicine and is also the Chief of the Section of General Neurology, Neurological Institute, Taipei Veterans General Hospital. She had won numerous awards and in 2010 was named as an outstanding physician by the Veteran's Affair Commission of Taiwan. She has published more than 200 papers in reputable journals and had been invited as a speaker at both national and international conferences. Her other professional activities include holding multiple positions as the Director of Taiwan Neurological Society, President of the Taiwan Dementia Society and supervisor of the Taiwan Headache Society.
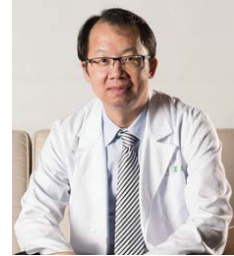
Shuu-Jiun Wang received an MD degree from the School of Medicine, National Yang-Ming University (NYMU), Taipei, Taiwan, in 1988. He is a practicing neurologist and Chair Professor and the Director of the Brain Research Center, NYMU, and the Director of the Neurological Institute, Taipei Veterans General Hospital (TVGH). He is also a headache specialist, and started the Headache Clinic in TVGH since 1996. He was the President of Taiwan Headache Society from 2005 to 2011. He is now the Presidents of the Taiwan Neurological Society and Asian Regional Consortium of Headache. Professor Wang is interested in clinical studies, epidemiology and neuroimaging studies of headache medicine, focusing on chronic migraine, cluster headache, reversible cerebral vasoconstriction syndrome, and spontaneous intracranial hypotension. He has authored over 300 scientific papers. He serves as an Associate Editor of Cephalalgia, Headache, the Journal of Headache and Pain, and other journals. He received the Bruce S. Schoenberg's International Neuroepidemiology Award from the American Academy of Neurology in 1994, the Seymour Solomon Lecture Award from the American Headache Society in 2010, and the G. Nappi Cluster Headache Award from the European Headache Society in 2016.