EEG analysis of seizure patterns using visibility graphs for detection of generalized seizures

Lei Wang a, c, *, Xi Long a, b, *, Johan B.A.M. Arends a, c, Ronald M. Aarts a, b

a Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands
b Philips Research, Eindhoven, The Netherlands
c Epilepsy Center Kempenhaeghe, Heeze, The Netherlands

HIGHLIGHTS

- A single-channel EEG signal is mapped into visibility graphs (VGS) so that network theory can be used.
- The power-law degree distributions in difference VG (DVG) show the best separation among the seizure EEG and non-seizure EEG.
- The connecting structure of horizontal VG (HVG) outperforms those of VG and DVG in distinguishing seizure EEG from non-seizure EEG.
- The proposed VGS-based features can help improve seizure detection for patients with intellectual disability.

ARTICLE INFO

Article history:
Received 16 January 2017
Received in revised form 5 June 2017
Accepted 13 July 2017
Available online 19 July 2017

Keywords:
Visibility graph (VG)
Horizontal VG (HVG)
Difference VG (DVG)
EEG seizure pattern
Seizure detection
Intellectual disability

ABSTRACT

Background: The traditional EEG features in the time and frequency domain show limited seizure detection performance in the epileptic population with intellectual disability (ID). In addition, the influence of EEG seizure patterns on detection performance was less studied.

New method: A single-channel EEG signal can be mapped into visibility graphs (VGS), including basic visibility graph (VG), horizontal VG (HVG), and difference VG (DVG). These graphs were used to characterize different EEG seizure patterns. To demonstrate its effectiveness in identifying EEG seizure patterns and detecting generalized seizures, EEG recordings of 615 h on one EEG channel from 29 epileptic patients with ID were analyzed.

Results: A novel feature set with discriminative power for seizure detection was obtained by using the VGS method. The degree distributions (DDs) of DVG can clearly distinguish EEG of each seizure pattern. The degree entropy and power-law degree power in DVG were proposed here for the first time, and they show significant difference between seizure and non-seizure EEG. The connecting structure measured by HVG can better distinguish seizure EEG from background than those by VG and DVG.

Comparison with existing method: A traditional EEG feature set based on frequency analysis was used here as a benchmark feature set. With a support vector machine (SVM) classifier, the seizure detection performance of the benchmark feature set (sensitivity of 24%, FD/h of 1.8 s) can be improved by combining our proposed VGS features extracted from one EEG channel (sensitivity of 38%, FD/h of 1.4 s).

Conclusions: The proposed VGS-based features can help improve seizure detection for ID patients.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Electroencephalography (EEG) is considered an important biometric for the diagnosis and screening of epileptic seizure detection in clinical practice. Many EEG-based seizure detection methods have been reported, aiming at differentiating non-seizure and seizure states for epileptic patients by using certain features extracted from EEG (Acharya et al., 2013; Alexandros et al., 2012; Ramgopal et al., 2014). These EEG features can be analyzed in different domains. For example, in the time domain, EEG amplitude, shape, and duration (Gotman et al., 1979), chaoticity or predictability measured by Lyapunov exponents (Stam, 2005), complexity evaluated by Lempel–Ziv complexity, as well as regularity by means of entropy measures (Acharya et al., 2012) have been studied. In the frequency domain, previous works have been addressing on characterizing epileptic seizures by means of EEG spectrum analysis.

* Corresponding authors at: Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands.
E-mail addresses: Lei.Wang@TUE.nl (L. Wang), x.long@tue.nl (X. Long).
In addition, some methods have been applied to analyze EEG for seizure detection in the time-frequency domain (e.g. discrete wavelet transform (Saab and Gotman, 2005) and empirical mode decomposition (Aloaiby et al., 2014)) as well as in the spatial-temporal domain (e.g. phase locking synchrony (Mormann et al., 2000)).

In this work, we propose to analyze EEG signals in the ‘visibility’ domain, in which visibility algorithms can be used to characterize a time series signal. The concept of using the visibility algorithms to analyze time series was proposed by Lacasa et al. (2008). The visibility algorithm can be interpreted as a geometric transform of a periodic time series, which can be an analogy of the discrete Fourier transform (DFT) in the frequency domain. The visibility algorithms can capture the presence of nonlinear correlations of a time series such as chaotic behavior where DFT fails to capture (Luque et al., 2009). Among the visibility algorithms, a visibility graph (VG) method has been previously used to characterize one-dimensional signal by mapping it into a two-dimensional network in a graph so that signal properties are geometrically visible (Lacasa et al., 2008). The VG methods have been applied to analyze electrophysiological signals such as ECG (Long et al., 2014) and EEG signals (Zhu et al., 2014), to exploit human sleep stages. For analyzing epileptic seizures, VG-based methods have been used for seizure detection based on the high-frequency sub-band of electrocorticography (ECoG) signals (Tang et al., 2013). The ictal and interictal EEG may be identified with the power law characteristics based on a VG method (Reijniesd et al., 2007). We thus evaluated the VG method in this study. In addition, we investigated the horizontal VG (HVG) (Luque et al., 2009), and difference VG (DVG) (Zhu et al., 2014), which are the extended versions of the basic VG. They are expected to capture different properties of EEG signals, e.g., HVG can better represent chaotic characteristics of EEG than VG (Luque et al., 2009), and DVG can help remove redundant information in VG (Zhu et al., 2014). The VG, HVG, and DVG are termed as VGM.

The seizure detection (Wang et al., 2015, 2016, 2017) for a specific population with both epilepsy and intellectual disability (ID) is challenging based on existing EEG features due to the presence of abnormal EEG activities caused by cerebral development disorders (Steffenburg et al., 1998; Guerrini et al., 2001). Clinicians often encountered different types of EEG in ID patients, such as abnormal background EEG (slow activity, no alpha), frequent occurrence of focal anomalies, high levels of inter-ictal epileptic transients that resemble seizures, abnormal sleep and wake cycles (difficult to interpret sleep/drowsiness EEG), as well as different seizure discharge patterns from non-ID epileptic patients. These EEG abnormalities may affect the seizure detection performance. It is because that seizure detection is to detect the abnormal ictal EEG, while the increased amount of abnormal interictal EEG causes more interference to the task of seizure detection. For example, an association has been found between intractability and abnormal EEG background in childhood epilepsy (Co and Holmes, 1999). The VGS-based features can provide supplementary ‘visibility’ information to the existing EEG features in the conventional time and frequency domain. Therefore, they might potentially benefit the seizure detection in ID patients. In addition, EEG signals vary across individuals, which often limits the application of a seizure detector in a large population. The variance of EEG signals in ID patients is often larger than that in non-ID patients, due to different levels of brain development. Our current study suggests that the ID population is a heterogeneous entity, causing significant variance of detection performance across subjects. However, when a one-dimensional EEG signal is mapped into a two-dimensional VGS, the geometrical properties of VGS are more robust to the variance in the original time and frequency space. For example, a mapped VG network can remain less variant when the original EEG signal has higher variances such as horizontal rescaling (i.e., frequency), vertical rescaling (changed amplitude) and addition of a linear trend (baseline drift) (Lacasa et al., 2008). This property of VGS-based methods is expected to reduce the variability of extracted EEG features within and across subjects. We thus evaluate the feasibility of the VGS-based features on seizure detection in the ID population.

Most of previous automatic seizure detection studies consider seizure detection as a classification task between the seizure class and non-seizure class. However, within the seizure class, the ictal EEG generally shows various morphological patterns. We term such ictal morphological patterns as the EEG seizure patterns (or seizure patterns), also known as epileptiform discharges (Ko, 2016) or polymorphic seizure patterns (Meier et al., 2008). Note that EEG seizure patterns differ from the clinical seizure types defined by ILAE (Fisher et al., 2017), which are diagnosed based on both EEG and clinical symptoms. The seizure patterns have been shown to associate with seizure detection performance (Wang et al., 2017; Meier et al., 2008). Typical seizure patterns include spikes, spike-wave complexes, and sharp/slow waves (De Lucia et al., 2008). In addition, the seizures accompanied by electromyography (EMG) artifacts are common in this ID population, and it is not appropriate to simply exclude them as artifacts (Conradsen et al., 2011). Therefore, we define them as the EMG seizures (Wang et al., 2017).

Fig. 1 illustrates the four typical seizure patterns in the ID patients. These patterns associate with clinical seizure types defined by ILAE. Specifically, fast spikes exist in most tonic seizures; spike-wave patterns often occur during absence-like seizures or at the end of tonic–clonic seizures; slow waves may present during focal seizures and rhythmic delta/theta seizures; and seizure-related EMG can exist in most tonic, tonic–clonic and myoclonic seizures. The spike-waves and rhythmic waves can also occur in interictal EEG with shorter durations (Wang et al., 2017). The four seizure patterns may occur in a sequence or as a combination (e.g., polyspike complexes) during a seizure or non-seizure EEG. The EEG analysis in the visibility domain may enlighten its applicability in the recognition of seizure patterns, thereby facilitating the diagnosis of clinical seizure types.

2. Materials and methods

2.1. Subjects and EEG signals

This was a retrospective study. We collected data of 29 adult epileptic patients (12 females, age 29±13) with an intellectual disability (3 light, 11 moderate, 15 severe, with IQ range at light [50–70], moderate [30–50], and severe [0–30]) from the data
archive in the Epilepsy Center Kempenhaeghe. The detailed subject demographics can be found in our previous study (Wang et al., 2017). For each seizure pattern, at least 10 patients were selected. Note that for some patients, their EEG can show more than one seizure pattern. For each patient, continuous scalp EEG signals of one day were acquired using 24 electrodes of Ag/AgCl in positions according to the 10–20 positioning system (Vázquez et al., 2012) with a sampling rate of 100 Hz. It is an ambulant recording measured by the EEG recording equipment TMS (Twente Medical Systems) and the software system BrainRT™. Each patient has at least one seizure recorded in the obtained EEG recordings. This EEG dataset has total EEG recordings of 615 h, and it contains 91 seizure events (89 generalized seizures, 2 focal seizures), including 21 tonic-clonic seizures, 16 tonic seizures, 9 myoclonic seizures and other unclassifiable seizure types (Fisher et al., 2017). The accumulated seizure duration across the 29 subjects is 3034 s. Two focal seizures presenting in this dataset were excluded for this study. Moreover, for this ID population, the generalized seizures account for more than 97% of the total number of seizures and focal seizures only 3% in real life (Nijsen et al., 2005).

### 2.2. EEG signal preprocessing

In this selected dataset, only two focal seizure events appeared and they were excluded. The remaining 89 seizures were all generalized seizures. Therefore, each single EEG channel contained seizure information in each seizure event. Here one EEG channel in the middle of scalp Cz from a unipolar montage was chosen for each subject. The EEG signals on Cz were then filtered by using a 10th-order Butterworth bandpass filter with the lower and the higher cutoff frequency of 0.5 Hz and 45 Hz, respectively (Diez et al., 2008). The filtered EEG signal was split into non-overlapping windows of two seconds. We keep only the EEG epochs with the amplitude range (i.e., a half of peak-to-peak amplitude) between 10 μV and 200 μV for further analysis. The lower bound (10 μV) was to reject the artifacts caused by loose electrode-skin contact or sweating (Uriguen and Garcia-Zapirain, 2015). The upper bound (200 μV) was to reject excessive artifacts caused by movements, and excessive EMG activities.

### 2.3. Annotations

The epileptic seizure annotation was conducted by EEG technicians based on visual inspection and was controlled by a neurologist specialized in epilepsy EEG signals. The seizure annotations included the clinical seizures types according to the definition of ILAE, the onset and offset of a seizure event and the epoch-by-epoch seizure patterns as shown in Fig. 1. Note that apart from the four predefined seizure patterns, there are also mixed (i.e., combination of multiple seizure patterns) and unclassifiable seizure patterns. The continuous EEG recordings on one single EEG channel, Cz from all 29 patients were collected, as shown in Table 1. Note that the partition in Table 1 aimed to evaluate the discriminative power of proposed EEG features on EEG epochs with different seizure patterns. We did not partition the EEG signals based on the clinical seizure types, because each clinical seizure type, e.g., tonic seizure, may contain several different EEG seizure patterns.

### Table 1

<table>
<thead>
<tr>
<th>EEG categories</th>
<th>Non-seizure</th>
<th>Seizure†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of EEG Epochs (2-s)</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>1106622</td>
<td>121</td>
<td>56</td>
</tr>
</tbody>
</table>

† It includes the seizure patterns: A (fast spike), B (spike-wave), C (wave), D (seizure-related EMG), M (mixed and unclassifiable seizure patterns).

Moreover, the diagnosis of clinical seizure types also depends on patients’ non-EEG symptoms.

### 2.4. Visibility graphs (VGS)

We briefly introduce the visibility algorithms that generate VGS, including the basic visibility graph (VG), horizontal VG (HVG) and difference VG (DVG).

#### 2.4.1. Visibility graph (VG)

A time series can be mapped into a VG network. To formalize this process, assuming a time series with n points, {x_i}, i = 1, ..., n, we denote each point as a node v(t, x_i) that corresponds to a vertical bar in a two-dimensional space. The location of the vertical bar is t, and the height of the bar is x_i. If the tip of the vertical bar can be connected using a straight line without obstruction to a tip of another vertical bar, then the corresponding two nodes in a VG network are connected by an undirected edge. That is, two nodes v(t, x_i) and v(t, x_j) in VG are connected if and only if the following rule is satisfied (Lacasa et al., 2008):

$$\forall k \in (i, j), \quad \frac{x_j - x_k}{j - i} = \frac{x_j - x_i}{j - i}.$$  

where i or j denotes the location of a point in a time series, x_i or x_j denotes the point value. Fig. 2 shows the process that a random time series with seven points is mapped into a VG network. A VG network can be noted as G_{VG}(V, E), where V and E are the nodes and edges of the graph, respectively. The degree δ(i) of i-th node v_i is the number of connected edges on this node.

#### 2.4.2. Horizontal VG (HVG)

It is known that EEG signals demonstrate chaotic behaviors (Korn and Faure, 2003). As an extended version of VG, HVG can represent such chaotic characteristics and easily discriminate chaos from uncorrelated randomness (Luque et al., 2009). In a HVG network, two nodes v(i, x_i) and v(j, x_j) are connected if and only if:

$$\forall k \in (i, j), x_k < x_i \quad \text{and} \quad x_k < x_j,$$

where i or j denotes the location of a point in a time series, x_i or x_j denotes the point value. Fig. 3 shows the process that a random time series with seven points is mapped into an HVG network. Two vertical bars (with height x_i and x_j) are connected when any bars between them are shorter than x_i and x_j. Accordingly, the two nodes in HVG network are connected by an undirected edge. Every two neighbor nodes are connected (as shown in Fig. 3). A HVG network can be noted as G_{hvg}(V, E), where V and E are the nodes and edges of...
the graph, respectively. The degree $\delta(i)$ of i-th node $v_i$ is the number of connected edges on this node.

### 2.4.3. Difference VG (DVG)

Assuming $G_{vg}(V, E_1)$ and $G_{vg}(V, E_2)$ be a VG and a HVG network respectively mapped from the same time series $\{x_t\}_{t=1, \ldots, n}$. The nodes set $V$ is the same for both the VG and HVG network. $E_1$ and $E_2$ are the edge sets of the VG and HVG, respectively. $G_{vg}(V, E_1)$ of DVG is defined by $E_2 = E_1 - E_2$. The degree sequence of DVG satisfies $\delta_{vg}(i) = \delta_g(i) - \delta_{bg}(i)$, where $\delta_{vg}(i), \delta_g(i)$, and $\delta_{bg}(i)$ are the degrees of node $v_i$ of the DVG, VG, and HVG network, respectively. DVG network could be beneficial to obtain the essential features of input signals (Zhu et al., 2014). Since DVG was obtained by subtracting the HVG from VG, the local neighborhoods of connected nodes, e.g., many nodes with degree value $\delta = 2$ on VG and HVG, can be removed in DVG (Donner and Donges, 2012), thereby only the edges of nodes that correspond to the peaks (or valley) of original EEG sequences are kept. Such nodes correspond to the important spikes that are morphological characteristics of seizure EEG signals.

![Figure 3](image)  
**Fig. 3.** An example of HVG network mapped from the same random time series used in Fig. 2. The mapped VG network is shown in the bottom with the degree sequence $\delta$: (1, 2, 4, 3, 2, 3).  

### 2.5. Characteristics of VGS

#### 2.5.1. Mean degree (MD)

Mean degree (MD) is the average value of degree $\delta_i$ of each node in a network. MD of VG, HVG, and DVG are denoted as $\text{MD}_g$, $\text{MD}_{hvg}$, and $\text{MD}_{bg}$, respectively. Note that MD of DVG is given by $\text{MD}_{bg} = \text{MD}_g - \text{MD}_{hvg}$. MD is mathematically equivalent to a commonly-used network metric, average clustering coefficient (ACC) (Watts and Strogatz, 1998). ACC measures the network's potential modularity (Ravasz et al., 2002), i.e., the likelihood that neighbors of a node are connected to each other. Therefore, instead of ACC, only MD was investigated in this work. Note that the sum of degree $\delta$ on all nodes is two times of the total number of edges $E$ in a network (i.e., $\text{MD} \propto E$). The value of MD thus also determines the entropy of network ensemble (Bogacz et al., 2006).

### 2.5.2. Degree entropy (DE)

The degree distribution (DD) in a network is the probability distribution that a node has a degree of $\delta$ (Zhu et al., 2014). The probability $P_{vg}(\delta)$ is computed as the number of nodes having degree $\delta$ divided by the total number of nodes in a VG network. We use $P_{vg}(\delta)$, $P_{hvg}(\delta)$ and $P_{bg}(\delta)$ to denote the DD of VG, HVG and DVG, respectively. Note that $P_{bg}(\delta) = P_{vg}(\delta) - P_{hvg}(\delta)$. In addition, the Shannon entropy of a degree in a network is formulated as,

$$H_G = -\sum_{\delta_i \in S} P_G(\delta_i) \log_2(P_G(\delta_i)),$$

where $G$ denotes a network, e.g., VG. $\delta_i$ is the i-th value of degrees in G. $S$ is the set of all possible values of degree $\delta_i$. $P_G(\delta_i)$ is the probability of $\delta_i$. Applying this formula on the networks VG, HVG and DVG, respectively, we obtain their DE: $H_{hvg}$, $H_{hvg}$, and $H_{bg}$.

#### 2.5.3. Power-law degree power (DP)

The degree distributions often have a power-law tail in a scale-free network (Albert and Barabási, 2002), or at least over some significant range of degree (Clauset et al., 2004), such that $P(\delta) \sim \delta^{-\lambda}$, $\delta_1 \leq \delta \leq \delta_2$, where $\delta_1$ and $\delta_2$ denote a range of degree, $\lambda$ is a constant, named degree power. The value of DP $\lambda$ is related to the fractality of a network (Lacasa et al., 2008).

#### 2.5.4. Assortativity coefficient (AC)

The assortativity coefficient (AC) can measure the overall connecting structure of a network (Newman, 2002). For a node in a network, it takes the preference of its connections to high- or low-degree nodes into account. Considering a network including a total of M edges, the i-th edge connects two nodes with degree of $\alpha_i$ and $\beta_i$ at their ends. The AC of this network is calculated as

$$\zeta = \frac{M^{-1} \sum_i \alpha_i \beta_i - [M^{-1} \sum_i \frac{1}{2} (\alpha_i + \beta_i)]^2}{M^{-1} \sum_i \frac{1}{2} (\alpha_i^2 + \beta_i^2) - [M^{-1} \sum_i \frac{1}{2} (\alpha_i + \beta_i)]^2},$$

where value of $\zeta$ is between $-1$ and $1$. If $\zeta > 0$, the network is assortative, in which case the high-degree (or low-degree) nodes are more likely to be connected to each other than to the low-degree (or high-degree) nodes; if $\zeta < 0$, the network is randomly mixed; and if $\zeta = 0$, the network shows disassortativity, in which case the high-degree nodes and the low-degree ones tend to mutually connect (Long et al., 2014). For example, in Fig. 4, AC of pattern C, $\zeta_C = 0.03$. It shows the pattern C is assortative. Patterns A, B, and D show disassortativity with $\zeta_A = -0.06$, $\zeta_B = -0.4$, and $\zeta_D = -0.3$. The pattern B and D have more high-degree nodes connecting with low-degree nodes than that in pattern A.

#### 2.5.5. Average shortest path length (SL)

The average shortest path length, SL describes global interconnectedness of a network (Smit et al., 2008). It is a value indicating the average number of steps required to go from each node to all others taking the shortest connected edges in a network. Watts and Strogatz (Watts and Strogatz, 1998) showed that ordered graphs (regular networks) are characterized by larger SL while random graphs have shorter SL. A network with small SL and high average clustering coefficient tend to be a so-called 'small-world' network (Smit et al., 2008). We compute SL on VG (SL$_{vg}$) and HVG (SL$_{hvg}$), in which each node is connected to its neighbor nodes with at least one edge. Note that in DVG, there are isolate nodes, i.e., nodes not connected with any other nodes. Therefore, SL is not computed on DVG.

#### 2.6. Statistical analysis

We computed a set of VGS-based features for each 2-s EEG epoch. It includes $|\text{MD}_g|$, $\text{AC}_g, \text{H}_{hvg}, \text{MD}_{hvg}$, $\text{MD}_{bg}$.
ACavg, Havg, Savg, SLavg, λ1, λ2), where λ1 and λ2 is the DP on certain degree intervals (refer to Section 3.1). These features represent certain properties of brain activities. To show their separability between non-seizure EEG and EEG with different seizure patterns, we measured the statistical significance by using a two-tailed Mann–Whitney U test between each two EEG categories (as described in Table 1), e.g., EEG with spike-wave pattern and non-seizure EEG. The features showing significant difference between non-seizure and seizure EEG were selected as the feature set for seizure detection.

2.7. Seizure detection

2.7.1. Cross validation (CV)

To evaluate the seizure detection performance, a support vector machine (SVM) classifier with a Gaussian kernel was employed to classify seizure and non-seizure EEG epochs. The optimal parameters of SVM were determined by using a grid search (Wang et al., 2017). Due to the limited duration of seizure EEG in this dataset, we used 5-fold CV to represent actual detection performance. That is, each seizure and non-seizure class was randomly partitioned into five equal parts, with four parts as a training set, remaining one part as a testing set. We average the five results on different testing sets as the final detection performance.

2.7.2. Performance criteria

The precision and recall (P-R) curve (He and Garcia, 2009), a plot of the sensitivity (i.e., #epoch of true positives/#real seizure epochs) vs. positive predictive value (PPV) (i.e., #epoch of true positives/#epochs of positive predictions), is known as a more suitable performance metric than the receiver operating characteristic (ROC) curve (Mormann et al., 2005) (i.e., a plot of the sensitivity vs. 1–specificity) in unbalanced datasets (Lpez et al., 2013). The area under curve of P-R curve (AUCPR) is a more discriminative performance indicator than AUCROC (Wang et al., 2017). The F1 score (Sun et al., 2009) (or F-measure) can be used to determine an optimal trade-off between sensitivity and PPV, achieving a desirable clinical detection performance. To quantify the discriminative power of proposed EEG features in terms of classification performance, we used the accumulated time (sec) of false detection (FD) per hour of EEG recordings, termed as FD/h. Compared with a more clinical indicator, number of FD per hour (FD/h), FD/h is more accurate in terms of qualifying the length of (miss-)agreement between classification result and human experts’ annotation in the non-seizure part, while FD/h reports only the accumulated number of misagreement. To provide comparable results with other EEG studies, we also used the performance criteria: the classification accuracy (ACC), the Cohen’s Kappa coefficient of agreement (Kappa). Details of these performance criteria can be found in Wang et al. (2017).

2.7.3. Comparison of feature sets

The EEG-based study in ID population is rare and there has been no an optimal EEG feature set for this population by far. Saab and Gotman (2005) reported a desirable seizure detection performance on a long-term scalp EEG dataset (non-ID population) by using EEG features based on frequency analysis. However, the reported performances in different populations are often not directly comparable due to substantially different EEG signals. We thus reproduce that feature set (as a benchmark feature set) and use it in our dataset to report a benchmark performance. This benchmark feature set contains six EEG features based on wavelet analysis, including the relative average amplitude, relative scale energy, and variation of amplitude, as well as the energy ratio of EMG activities (termed as EMG Amp Ratio in Saab and Gotman, 2005). The features are computed on each EEG channel and we average the obtained features on all 21 EEG channels. Note that three EEG electrodes above eyes including Fp1, Fpz, and Fp2 are excluded to avoid eye movement artifacts. To evaluate the seizure detection performance of the VCS-based features, we compare it with the benchmark feature set on the same dataset. In addition,
the usefulness of VGS-based features is tested by combining them with the benchmark feature set. That is, for each EEG epoch in a 2-s sliding window, the extracted VGS-based features and the benchmark features are assembled as a combined feature set, resulting a total of 17 features in the combined feature set.

3. Results

3.1. Comparison of degree distributions (DDs)

Each EEG epoch in a 2-s sliding window without overlap was mapped to networks, VG and DVG. To show the overall DDs of a network, we pooled over all 2-s EEG epochs in each category, then computed the probability of a degree, $p(\delta)$ over all ‘pool-over’ degrees. The pool-over DDs of VG and DVG are shown in Fig. 5a and b, respectively. However, for an on-line analysis of EEG signals, a DD can be computed only on a current EEG epoch. To this end, we computed the average DDs across all epochs in an ‘epoch-by-epoch’ manner. In which, the DDs, $P_{\text{avg}}(\delta)$ and $P_{\text{avg}}(\delta)$ average the probability $p(\delta)$ of a degree $\delta$ across all epochs in a EEG category, as shown in Fig. 5c and d, respectively.

Fig. 5 shows that DVG can better discriminate the differences of DDs across EEG categories, compared with VG (also HVG that is not shown here). The pool-over DDs on all EEG categories show the power-law tails on both VG and DVG networks. However, the DDs on the lower interval of degrees $\delta$ are different on VG and DVG networks. One possible explanation is that by subtracting the degrees of HVG from VG, the connections of local neighborhoods in VG which often correspond to low-degree nodes (mainly represented by HVG) are removed. Therefore, it makes the proportion of nodes with higher degrees in DVG larger, so called ‘the rich getting richer’. Accordingly, a DVG tends to be a scale-free network (Reijneveld et al., 2007), and it thus shows a global property of power-law topology.

The pool-over DDs of DVG show better separation than that on VG in the lower interval of degrees. Meanwhile, the degrees in lower interval have higher probability of occurrence because of higher value of $P(\delta)$. The DDs at the lower degree interval therefore can be used for discriminating EEG with seizure patterns from the non-seizure EEG (as shown in Fig. 5d). The DD in each category approximately complies with the power-law topology, i.e., $P_{\text{avg}}(\delta) \sim \delta^{-\lambda}$ in two degree intervals. Therefore, we propose to use the DP (i.e., the minus slope value in a log-log plot): $\lambda_1$ and $\lambda_2$ on
the degree intervals [1–10] and [10–25], respectively. The DP $\lambda_1$ and $\lambda_2$ can be used for detecting seizure EEG epochs and further distinct seizure patterns. For example, $\lambda_1$ on the EEG categories, ‘0’, ‘A’, ‘B’, ‘C’, and ‘D’ are 0.5, 0.8, 0.0, 0.3, and 1.2, respectively. The overlapped DDs at larger degree $\delta$ may be due to the outliers in each seizure pattern caused by the inaccurate labeling in the process of visual inspection-based annotation. However, the impact of such outlier degrees is small because of the low value of probability $P(\delta)$. The epoch-by-epoch DDs have more outliers because of small sample size in each 2-s epoch. Therefore, compared with the pool-over manner, the DDs of the epoch-by-epoch DVG tend to be ‘blurry’ at larger degree $\delta$.

3.2. Comparison of VGS characteristics

The HVG was used to obtain DVG so that the EEG ‘spikes’ (i.e., nodes with high degrees in network) in DVG are highlighted from the background EEG signals. Therefore, the features extracted from DVG are expected to be more discriminative than that on VG. The MD and DE can characterize the degree-related properties of networks. In Fig. 6, the features $M_{DVG}$ (right in row 1) and $H_{DVG}$ (right in row 2) can clearly separate non-seizure EEG and EEG with different seizure patterns (except the pair denoted with asterisk). The statistical significance on VG and DVG are similar, and they are better than VG. The $A_{DVG}$ (middle in row 3) can separate the EEG categories better than that on VG and DVG, which indicates that apart from degree-related characteristics, the connecting structure of networks could be another important characteristic for identifying seizure EEG.

The SL measures the overall connecting strength of a network. The more high-degree nodes in a network, the shorter for each node to reach other nodes. Fig. 7 shows that $S_{VHG}$ (left) can better separate the EEG categories than $S_{HVG}$. In addition, $S_{HG}$ has general smaller value, indicating that VG is a ‘small-world’ network (given that VG also has high MD as shown in Fig. 6). For example, the spike-wave EEG (pattern B in left of Fig. 7), has the smallest values of SL. It is because the peaks of the spikes from the periodic spike-wave EEG signals are mapped to the high-clustering nodes in a network that shorten the path between any two nodes.

Fig. 8 shows that the DP $\lambda_1$ and $\lambda_2$ show statistical significance ($p<0.0001$) in all pairs of EEG categories except one pair on $\lambda_2$. The $\lambda_1$ has less outliers than $\lambda_2$, and it clearly discriminates different EEG categories. The estimation of the DP depends on a finite number of non-null degree values. As the lower-value degrees $\delta$ have higher probability $P(\delta)$ (as shown in Fig. 5d), thus it is more likely to have enough number of non-null degrees in an EEG epoch to estimate an accurate value for $\lambda_1$. The DP $\lambda_1$ is therefore the most robust one in the VGS-based features.

3.3. Seizure detection performance

The DVG is generated from networks VG, thus redundant information may be contained in the features extracted from VG and DVG networks. The Spearman’s rank correlation coefficient $r$ between each pair of features from VG and DVG is found to be high, $[r = 1.000, 0.900, 0.994, p < 0.00001]$ on feature pairs of ($M_{VG}$, $M_{DVG}$), ($A_{VG}$, $A_{DVG}$), and ($H_{VG}$, $H_{DVG}$), respectively. Therefore, to avoid using redundant features, we employ only the DVG-based features: ($M_{DVG}$, $A_{DVG}$, $H_{DVG}$). Combining with HVG-based features: ($M_{HVG}$, $A_{HVG}$, $H_{HVG}$), SL and DP, we obtain the final VGS feature set: ($M_{HVG}, M_{DVG}$, $H_{HVG}$, $H_{DVG}$, $A_{HVG}$, $A_{DVG}$, $A_{SL}$, $S_{HVG}$, $\lambda_1$, $\lambda_2$).

The performance of seizure detection by using SVM is shown in Fig. 9, in which average P-R curves on 5-fold testing sets on the VGS benchmark and combined feature set are plotted. The optimal trade-off (markers on the curves) between sensitivity and PPV on each P-R curve is determined by maximizing $F_1$ score. The detection performance at corresponding marker is shown in Table 2. Although the feature set of VGS (‘VG’ in Fig. 9) alone shows low performance, it can help improve the performance of the benchmark method (‘Benchmark’ in Fig. 9). That is, the combined feature set (‘VG + Benchmark’ in Fig. 9) achieves the best detection performance, with epoch-based sensitivity of 0.38, average FD/h of 1.4s. Note that VGS features were computed from only one EEG channel, while the benchmark features were computed from 21 EEG channels and used the average value of each feature extracted from all valid EEG channels after preprocessing. Although one single EEG channel is easier to obtain and less invasive than using multi-channels, it may easily suffer from the interruption of presence of interictal epileptiform discharges. This is one reason why the performance of VGS feature set is low. In our future work, we will compute the VGS features on all valid EEG channels (with high computational cost), and compare it with the performance by using one single channel.

4. Discussion

For EEG-based seizure detection in ID patients, the EEG analysis from the transformed complex networks provides a new ‘visibility’ domain other than the traditional time and frequency domains. In the visibility domain, the chaotic and fractal behaviors of EEG signals can be characterized and used for distinguishing seizure patterns. The statistical values of the average SL and MD show the VG tend to be a ‘small-world’ network. The DVG shows typical power-law topology on the DD-dependent intervals. It suggests that the DVG mapped from EEG signals may be a scale-free network. Applying the associated properties of scale-free networks, we can extract features such as $M_{DVG}$, $H_{DVG}$, and DP (i.e., $\lambda_1$, $\lambda_2$) to discriminate EEG with seizure patterns from non-seizure signals. In addition, the analysis of DDs on VGS should be based on EEG of each seizure pattern because each pattern has a unique feature distribution (e.g., DD) that can be distinguished from others. Simply pooling over seizure EEG epochs with all different seizure patterns as one ‘mixed’ seizure class would possibly lead to no or small difference in feature values between seizure and non-seizure classes.

Surprisingly, the connecting structures of HVG can better discriminate EEG with different seizure patterns from non-seizure EEG signals than those on VG and DVG. It indicates that the connecting structure of network could be another important property (apart from DDs-related properties) for identifying seizure EEG signals. In addition, Luque et al. (2009) demonstrated that the network HVG is associated with the topological properties of chaotic series, and can clearly distinguish noise and high- (or low-) dimensional chaos systems. It suggests that the seizure EEG signals with different morphological patterns may represent chaos systems with different topological properties caused by seizures.

The performance of seizure detection based on traditional EEG features can be significantly improved by combining our proposed VGS features. However, the achieved performance may be still not desirable for automatic seizure detection in real clinical applications. The low performance mainly results from the intractability of this target ID population (e.g., one of six seizure events on average was clinically observed by specialized epilepsy nurses (Nijssen et al., 2005)). Specifically, these generalized seizures on this dataset have short duration (e.g., a half of seizures are shorter than 20 s). These minor seizures often show ambiguous EEG changes and unclear seizure boundaries (i.e., borders of ictal and interictal EEG are indistinct) (Fisher and Scharfman et al., 2017; Krauss, 2011) without clinical signs. As a result, the corresponding seizure EEG are difficult to detect from background activities. In addition, in this study, to quantify the discriminative power (i.e., classification performance)
of proposed EEG features, we report the detection performance in an epoch-based way (Wang et al., 2017). The epoch-based performance is negatively affected by the presence of interictal epileptiform discharges (IEDs) (King et al., 1998). A high level of spike-wave IEDs have been observed only in this specific ID population by far. Such spike-wave IEDs can occur during more than...
Table 2
Seizure detection performance of each feature set.

<table>
<thead>
<tr>
<th>Feature set</th>
<th>AUCPR</th>
<th>F1 score</th>
<th>ACC</th>
<th>Kappa</th>
<th>Sensitivity (recall)</th>
<th>Specificity</th>
<th>PPV (precision)</th>
<th>FD/1h (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGS (#10)</td>
<td>0.0676</td>
<td>0.1440</td>
<td>0.9976</td>
<td>0.1429</td>
<td>0.9148</td>
<td>0.1483</td>
<td>1.4035</td>
<td></td>
</tr>
<tr>
<td>Benchmark (#7)</td>
<td>0.2264</td>
<td>0.3002</td>
<td>0.9984</td>
<td>0.2995</td>
<td>0.9995</td>
<td>0.4010</td>
<td>1.7612</td>
<td></td>
</tr>
<tr>
<td>VGS + Benchmark (#17)</td>
<td>0.4087</td>
<td>0.4575</td>
<td>0.9987</td>
<td>0.4569</td>
<td>0.9996</td>
<td>0.5749</td>
<td>1.4032</td>
<td></td>
</tr>
</tbody>
</table>

nNumber of EEG features in a feature set.

Fig. 7. Boxplots of features SL on networks, VG (left), and HVG (right) on EEG categories: Non (non-seizure class), A (fast spike), B (spike-wave), C (wave) and D (EMG). The Mann-Whitney test shows significant difference (p < 0.05) between all pairs of groups except the pairs denoted by asterisks (∗) indicates p > 0.05. The outliers of SL_{lg} [6–25], SL_{hvg} [25–68] are cut off for easy visualization.

Fig. 8. Boxplots of proposed features power λ1 and λ2 on DVG on EEG categories, Non (non-seizure class), A (fast spike), B (spike-wave), C (wave) and D (EMG). The Mann-Whitney test shows significant difference (p < 0.0001) between all pairs of groups except the pair denoted by asterisks (∗) indicates p > 0.05. The outliers [−2 to −6] of λ2 are cut off for easy visualization.

50% of the EEG recordings on some ID patients, while less than 1% in the non-ID patients. However, the manual annotation or automatic recognition of such IDEs are difficult. It thus challenges both the off-line analysis and on-line monitoring for the ID patients. The use of context information of seizures may help solve this problem (Xun et al., 2016). In our future work, we will use a multi-domain feature set and the context information including also non-EEG signals (de Vel et al., 2013) (e.g., heart rate) to perform an event-based (or clinically-oriented) seizure detection.

5. Conclusions

We mapped single-channel EEG signals into complex networks through visibility graphs including VG, HVG and DVG. The characteristics of these networks were studied and compared across non-seizure EEG and EEG with four typical seizure patterns in 29 ID patients. We showed that the DD of DVG can clearly distinguish EEG with each seizure pattern from non-seizure EEG. The degree-based characteristics, e.g., MD, DE and DP, can be used as EEG features to detect a seizure pattern. The connecting structure measured by HVG can better distinguish EEG of different seizure patterns from non-seizure EEG than those by VG and DVG. Finally, we proposed a novel VGS-based feature set and demonstrated that it can help improve the seizure detection for ID patients.

References
