Neural decoding of electrocorticographic signals using dynamic mode decomposition

Yoshiyuki Shiraishi1, Yoshinobu Kawahara2,3, Okito Yamashita2,3, Ryohei Fukuma1,5, Shota Yamamoto1, Youichi Saitoh2,6, Haruhiako Kishima5 and Takufumi Yanagisawa1,5,7

1 Osaka University, Institute for Advanced Co-Creation Studies, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan
2 RIKEN Center for Advanced Intelligence Project, Nihonbashi 1-chome Mitsui Building, 15th floor, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan
3 Institute of Mathematics for Industry, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan
4 Department of Computational Brain Imaging, ATR Neural Information Analysis Laboratories, 2-2-2 Hikaridai, Seika-cho, Kyoto 619-0288, Japan
5 Department of Neurosurgery, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan
6 Department of Neuromodulation and Neurosurgery, Osaka University Graduate School of Medicine, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan
7 Department of Neuroinformatics, ATR Computational Neuroscience Laboratories, 2-2-2 Hikaridai, Seika-cho, Kyoto 619-0288, Japan

E-mail: tyanagisawa@nsurg.med.osaka-u.ac.jp

Keywords: dynamic mode decomposition, electrocorticography, neural decoding, Grassmann kernel, principal angle

Supplementary material for this article is available online

Abstract

Objective. Brain–computer interfaces (BCIs) using electrocorticographic (ECoG) signals have been developed to restore the communication function of severely paralyzed patients. However, the limited amount of information derived from ECoG signals hinders their clinical applications. We aimed to develop a method to decode ECoG signals using spatiotemporal patterns characterizing movement types to increase the amount of information gained from these signals.

Approach. Previous studies have demonstrated that motor information could be decoded using powers of specific frequency bands of the ECoG signals estimated by fast Fourier transform (FFT) or wavelet analysis. However, because FFT is evaluated for each channel, the temporal and spatial patterns among channels are difficult to evaluate. Here, we used dynamic mode decomposition (DMD) to evaluate the spatiotemporal pattern of ECoG signals and evaluated the accuracy of motor decoding with the DMD modes. We used ECoG signals during three types of hand movements, which were recorded from 11 patients implanted with subdural electrodes. From the signals at the time of the movements, the modes and powers were evaluated by DMD and FFT and were decoded using support vector machine. We used the Grassmann kernel to evaluate the distance between modes estimated by DMD (DMD mode). In addition, we decoded the DMD modes, in which the phase components were shuffled, to compare the classification accuracy. Main results. The decoding accuracy using DMD modes was significantly better than that using FFT powers. The accuracy significantly decreased when the phases of the DMD mode were shuffled. Among the frequency bands, the DMD mode at approximately 100 Hz demonstrated the highest classification accuracy. Significance. DMD successfully captured the spatiotemporal patterns characterizing the movement types and contributed to improving the decoding accuracy. This method can be applied to improve BCIs to help severely paralyzed patients communicate.

1. Introduction

A brain–computer interface (BCI), which enables communication via brain activity, is the only way to express thoughts for severely paralyzed patients who are completely unable to move their body [1]. BCIs are in high demand for these patients [2], who choose to undergo invasive procedures involving the placement of electrodes on the surface [3] or inside the cortex [4, 5]. An electrocorticographic (ECoG) signal,
which is measured by a surface electrode, is a promising signal for use in BCIs [6] due to its stability [7] and less invasive nature compared to penetrating electrodes [8, 9]. An ECoG-BCI was successfully applied to a severely paralyzed patient with amyotrophic lateral sclerosis (ALS) and enabled the patient to stably communicate for months [10]. However, much less information is derived from an ECoG signal than from penetrating electrodes [11]. The patient with ALS was able to type faster using unit activity recorded by a penetrating electrode, at more than 30 characters per minute [11], than using ECoG signals [10]. Thus, for the clinical application of ECoG-BCIs, some novel technologies to extract information from ECoG signals are necessary.

To decode the ECoG signals, the oscillation component is considered to have important information, and various power spectral density (PSD) estimations [12–14] have been used as features for machine learning [15–19]. Among the methods to evaluate the PSD [18], fast Fourier transform (FFT) is often used for online decoding because of the low calculation cost [20]. FFT evaluates the powers of some frequency bands for each channel in a time series and for individual time points in space. However, because these features are independently evaluated, FFT cannot evaluate the patterns of ECoG signals in both space and time simultaneously. Therefore, a method to characterize the nonlinear dynamics as the linear summation of coherent patterns across both space and time (e.g. frequency and phase) is necessary.

Dynamic mode decomposition (DMD) has recently attracted attention as a way of obtaining modal representations of nonlinear dynamics from general multivariate time series data without explicit prior knowledge about the dynamics [21, 22]. DMD is closely related to spectral analysis of the Koopman operator (see supplementary information (available online at stacks. doi.org/JNE/17/036009/mmmedia)) and has been successfully applied for the extraction of spatiotemporal patterns, broadly in fluid mechanics [21–24], but rarely in neuroscience, e.g. ECoG and functional magnetic resonance imaging (fMRI) [25–28]. Although a previous study using human ECoG signals demonstrated that the DMD was able to capture some characteristic spatiotemporal patterns related to spindles [27], the spatiotemporal patterns were not used for neural decoding. It has not been revealed whether the spatial and temporal coherent patterns identified by DMD improve the accuracy of the neural decoding used in BCI technology.

A problem in the application of DMD for neural decoding is that DMD transforms the spatial and temporal signals into a set of vectors representing the characteristic spatial mode, which has a characteristic frequency in the time domain. Notably, because the characteristic frequency of the DMD mode is estimated from single trial data, the frequency is not fixed across trials and varies depending on the original signals. Therefore, to classify the signals by the vectors estimated by DMD, proper metrics are needed to evaluate how the sets of vectors are similar among trials. Here, we used the principal angle as a metric to evaluate the relationship among the sets of vectors for each trial [29]. A Grassmann kernel was used with support vector machine (SVM) to classify the set of vectors. In this study, we classified upper limb movements using the DMD components with SVM. The classification accuracy using the DMD component was compared with that using the FFT powers and that using the DMD components whose phases were shuffled.

2. Methods

2.1. Subjects

11 patients (four females, seven males; age range, 13–66 years) with subdural electrodes participated in this study. Eight patients were the same patients from our previous report [17]. Five patients had different degrees of motor dysfunction and sensory disturbances in their upper limbs due to strokes without damage to the sensorimotor cortex, previous surgery on the motor cortex, or brain tumor on the sensorimotor cortex (table 1). Six patients with epilepsy had no motor dysfunctions or sensory disturbances. All participants or their guardians gave written informed consent to participate in the study, which was approved by the ethics committee of Osaka University Hospital (No. 08061, No. 19257).

2.2. Movement tasks

Experiments were performed approximately 1 week after electrode placement. The patients performed three types of upper limb movements, selecting among grasping, pinching, hand opening, thumb flexion, and elbow flexion [17]. We presented the patients with three sound cues at 1 s intervals. The patients performed one of three prespecified movements after the third sound cue. The patient selected the movement type for each trial, so the order and number of exercises were random. The numbers of the selected movement types are shown in table 2. The patients were instructed to perform the selected movement once, immediately after the third cue, and to return their hands to the resting position, which was an immediate hand posture among the performed movements. For the resting position, patients were instructed to relax their hands or elbows with slightly flexed joints. Each type of movement was performed approximately 30 to 100 times. The movement instructions were delivered using a PC monitor controlled by ViSaGe (Cambridge Research System, Rochester, UK) placed in front of the patients.

2.3. ECoG recording and preprocessing

ECoG signals were recorded at 1 kHz by an EEG-1200 system (Nihon Koden, Tokyo, Japan). The subdural
Table 1. Clinical profiles.

| Patient no. | Age (yr)/sex | Diagnosis                        | Paresis in affected limb (MMT) | Sensation in affected limb |
|-------------|--------------|----------------------------------|-------------------------------|---------------------------|
| 1           | 65/M         | Ruptured spinal dAVF             | Spastic (4)                   | Hypoesthesia              |
| 2           | 64/M         | R thalamic hemorrhage            | Spastic (4)                   | Hypoesthesia              |
| 3           | 20/F         | L intractable epilepsy           | None (5)                      | Normal                    |
| 4           | 34/F         | R intractable epilepsy           | None (5)                      | Normal                    |
| 5           | 22/F         | R intractable epilepsy           | None (5)                      | Normal                    |
| 6           | 14/M         | R intractable epilepsy           | None (5)                      | Normal                    |
| 7           | 13/M         | L intractable epilepsy           | None (5)                      | Normal                    |
| 8           | 33/M         | R intractable epilepsy           | Slightly spastic (4)          | Hypoesthesia              |
| 9           | 26/M         | L brain tumor                    | None (5)                      | Normal                    |
| 10          | 66/F         | R subcortical infarction         | Spastic (4)                   | Hypoesthesia              |
| 11          | 49/M         | R putaminal hemorrhage           | Slightly spastic (5-)         | Hypoesthesia              |

dAVF, dural arteriovenous fistula; F, female; L, left; M, male; MMT, manual muscle test; R, right.

Table 2. Movement tasks and recording conditions.

| Patient no. | Pre-specified movement tasks (label 1, 2, 3) | Trial count (label 1) | Trial count (label 2) | Trial count (label 3) | Number of electrodes |
|-------------|----------------------------------------------|-----------------------|-----------------------|-----------------------|---------------------|
| 1           | Grasp, pinch, open                           | 42                    | 46                    | 32                    | 60                  |
| 2           | Grasp, pinch, open                           | 19                    | 25                    | 33                    | 60                  |
| 3           | Grasp, pinch, open                           | 44                    | 36                    | 37                    | 20                  |
| 4           | Grasp, pinch, open                           | 34                    | 44                    | 49                    | 30                  |
| 5           | Grasp, pinch, open                           | 42                    | 39                    | 39                    | 30                  |
| 6           | Grasp, pinch, open                           | 34                    | 20                    | 22                    | 15                  |
| 7           | Grasp, pinch, open                           | 37                    | 50                    | 36                    | 27                  |
| 8           | Rock, paper, scissors                        | 31                    | 30                    | 39                    | 49                  |
| 9           | Rock, paper, scissors                        | 20                    | 20                    | 20                    | 24                  |
| 10          | Grasp, thumb, elbow                          | 36                    | 40                    | 36                    | 20                  |
| 11          | Grasp, thumb, elbow                          | 71                    | 69                    | 39                    | 32                  |

Electrodes were placed on the sensorimotor cortex based on clinical necessity. For each patient, 15–60 electrodes were implanted. All signals were subtracted by the average of the signals of all electrodes at each time point, a well-known method for the calculation of a common average reference [30, 31]. An epileptologist identified the channels with severe noise, and these were excluded from the analysis. Three channels for patient 7 and one channel for patient 10 were removed ( supplementary figures 3, 5).

For the following analysis, we used the ECoG signals at 0.5 s after the third sound cue, Sexecute, and at 1.0 s before the first sound cue, Snorm (figure 1). Here, Sexecute and Snorm were composed of n channels for 500 and 1000 time points, respectively. Here, we used 1.0 s of ECoG signals for normalization (Snorm) to decrease the variability in Sexecute. Notably, we also used 0.5 s of ECoG signals before the first sound cue for the Snorm, but the classification accuracy using the FFT feature did not significantly change with the normalization of 0.5 s (supplementary information, table 1). All processes were implemented by MATLAB (R2015b).

2.4. FFT features for neural decoding

As a conventional approach for the neural decoding of ECoG signals, FFT was used to create the FFT features of the signals [32]. FFT was applied as 512 points and 1024 points for Sexecute (500 points) and Snorm (1000 points), respectively, by zero-padding with a hammering window. Snorm was only used to normalize Sexecute to reduce variation among trials. The power features of four frequency bands were obtained by averaging the estimated power of the four frequency bands: theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and high gamma (80–150 Hz). In addition to these powers, we also calculated the average 0 Hz power for Sexecute and Snorm. Finally, these five features of Sexecute were normalized by the mean power pnorm and standard deviations σnorm of the corresponding features derived from Snorm, to create the FFT features, which consisted of five features by n channels [17, 33]. The features were obtained by  \( \frac{p_{execute} - p_{norm}}{\sigma_{norm}} \) (pexecute is the power feature vector of Sexecute). Here, the FFT powers of Sexecute were normalized because normalized features were shown to be decoded with better accuracy than FFT powers of Sexecute without normalization [15].

2.4.1. Computation of the dynamic mode decomposition.

Assuming that the sensorimotor cortex during exercise is one dynamic system, the system can be described using its observed variable x (ECoG signals) as follows:

\[
\frac{dx}{dt} = Ax + Bu + w(t)
\]

where A is a constant matrix, B is a constant input matrix, w(t) is a white noise vector, and u is a control input vector.
\[
\frac{dx}{dt} = f(x; t; \mu),
\]
where \(t\) and \(\mu\) are the time and system parameters. Because the actual recorded signals are values at discrete times, \(x_k = x(k \Delta t)\) is delimited by \(\Delta t\), and the dynamic system is described as \(x_{k+1} = F(x_k)\). Although \(F\) is difficult to obtain analytically, we can obtain linearly approximated dynamics \(x_{k+1} = Ax_k\) from the recorded signals. The DMD can be used to estimate \(A\) of unknown dynamics.

First, we briefly explain how the DMD was evaluated for column vector \(x_k\), which is the observations across all electrodes at the \(k\)-th time sample \([27]\). Gathering measurements from \(m\) snapshots in time, we constructed two \(n \times (m-1)\) raw data matrices, \(X_1\) and \(X_2\):

\[
X_1 = \begin{bmatrix} x_1 & x_2 & \ldots & x_{m-1} \end{bmatrix} \\
X_2 = \begin{bmatrix} x_2 & x_3 & \ldots & x_m \end{bmatrix}
\]

Here, the DMD of the data matrix pair \(X_1\) and \(X_2\) was given by the eigendecomposition of the transition matrix \(A\), which is an \(n \times n\) matrix satisfying \(X_2 = AX_1\).

Using the singular value decomposition (SVD) of the data matrix \(X_1 = U \Sigma V^*\), \(A\) was approximated as \(A \approx X_2X_1^+ \triangleq X_2V\Sigma^{-1}U^*\) (+ is pseudoinverse matrix of \(X_1\)).

However, the eigenvalue decomposition of \(A\) requires a tremendous amount of calculation. Therefore, we used the orthogonal matrix \(U\) to make \(\tilde{A} \triangleq U^T \times A \times U = U^T \times X_2V\Sigma^{-1}U^*\), which has fewer dimensions than \(A\). That is, we assumed that there exists some low-dimensional spatial structure in \(X_1\). Using \(\tilde{A}\), the calculation for the eigenvalue decomposition of \(A\) was reduced. The eigenvalues \(\lambda\) of the eigenvalue decomposition \(\tilde{A}\) were obtained by \(\tilde{A}W = \Lambda W\), with the eigenvalue matrix \(\Lambda\) approximately equal to the eigenvalue matrix of \(A\).

In addition, for the exact DMD, \(\hat{A} \triangleq \sum_{i=1}^{m-1} \tilde{A} \Sigma^{-1}\) was set to obtain the eigenvector \(W\) by applying the eigenvalue decomposition \(\hat{A}W = \Lambda W\). Then, by putting \(W = \sum_{i=1}^{m-1} W_i\), the DMD mode was obtained as each row of \(\hat{A} \triangleq X_1V\Sigma^{-1}U\Lambda^{-1}\).

Here, \(\hat{A}\) is the diagonal matrix of eigenvalues \(\lambda_i\). Each eigenvalue \(\lambda_i\) is a DMD eigenvalue. Each column of \(\hat{A}\) is a DMD mode \(\varphi_i\), corresponding to the eigenvalue \(\lambda_i\): \(i = 1, \ldots, m - h - 1\). The absolute values and phases of \(\hat{A}\) are the decay rates and frequencies, respectively.

Finally, we were able to write an approximation of the observed data as a simple dynamic model \(\hat{X}(t)\):

\[
\hat{X}(t) = \hat{\Phi}\exp(\Omega t)z,
\]
where \(\Omega = \log(\Lambda)/\Delta t\), \(t\) is time, and \(z\) is a set of weights to match the first time point measured such that \(\hat{X}(1) = \hat{\Phi}z\). It should be noted that \(\hat{\Phi}\) and \(\Omega\) are sets of vectors with complex values (figure 2).

DMD was originally used to evaluate the \(X_1\) of \(n \gg m\). However, in our data, the ECoG signal has approximately 100 channels for 500 ms. Therefore, to calculate the DMD, we constructed a new augmented data matrix \(X_{1,\text{aug}}\) as follows:
\[ S_{\text{execute}} = \sum_j \Phi_j \times \exp(A_{\text{real}} \times t + i \times A_{\text{imagin}} \times t) \times z_0 \]

\[ j=1..300 \]

Figure 2. DMD of ECoG signals during movements. The ECoG signals of n channels \((S_{\text{execute}})\) were converted into the \(n \times 300\) matrix of magnitude and phases that represents the complex values of the DMD modes \(\Phi_j\). \(A_{\text{real}}\) and \(A_{\text{imagin}}\) represent the decay rates and frequencies.

\[ X_{1\text{aug}} = \begin{bmatrix} x_1 & x_2 & \ldots & x_{m-h} \\ x_2 & x_3 & \ldots & x_{m-h+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_h & x_{h+1} & \ldots & x_{m-1} \end{bmatrix} \]

Here, \(h\) is a minimum integer that satisfies \(h < \frac{m+1}{n+1}\). \(X_{1\text{aug}}\) is an array generated from \(X_1\) of \(m\) length by stacking them in the channel dimension by sliding the step one by one. We constructed \(X_{2\text{aug}}\) using the same procedure.

The DMD features were obtained from \(X_{1\text{aug}}\) and \(X_{2\text{aug}}\) by the method previously mentioned. Notably, we generated \(X_{1\text{aug}}\) and \(X_{2\text{aug}}\) using only \(S_{\text{execute}}\) without any normalization by \(S_{\text{norm}}\) to avoid damaging the spatial relations among the channels of \(S_{\text{execute}}\) by subtracting the averages and dividing the standard deviations of \(S_{\text{norm}}\) for each channel. In fact, the DMD mode \(\hat{\Phi}\) will be changed by the normalization of \(X_{1\text{aug}}\) and \(X_{2\text{aug}}\) by the same \(S_{\text{norm}}\).

Here, the number of channels, \(n\), and the size of the \(X_{1\text{aug}}\) were different among the subjects. Therefore, to equalize the number of DMD modes, we selected the 300 modes with the largest singular values. Because the DMD mode \(\varphi_{i\text{aug}}\) yielded from \(X_{1\text{aug}}\) and \(X_{2\text{aug}}\) was a vector of \(hn \times 1\), we used the first \(n\) vector of \(\varphi_{i\text{aug}}\) corresponding to \(\varphi_i\), which is the mode for the original data \(X_1\) and \(X_2\) for the decoding task. In addition, because the size of \(\varphi_i\) generally decreases with the higher frequency range (imagery part of the eigenvalue \(\lambda_i\)), we normalized each mode by the L2 norm of the mode \(\varphi_i\). Notably, the principal angle requires the orthonormal matrix to be equalized among the permutation of the composed vectors. To simplify the calculation, we aligned the mode magnitude by the L2 norm instead of the orthogonalization.

2.4.2. Neural decoding by the Grassmann kernel.

The FFT feature is represented as a set of vectors of \(n \times 5\). The DMD features are composed of eigen-vectors \(\Phi_i\) corresponding to a characteristic frequency \(\omega_j\) (\(j = 1, \ldots, 300\)).

Therefore, the DMD feature is represented as a set of vectors \(n \times 300\). Notably, the DMD feature for classification is a subspace of the representation. To compare the subspaces for the classification task, we evaluated a principal angle, which measures the distance between the subspaces based on the set of vectors. Here, based on the principal angle, we used the Grassmann kernel to classify the set of vectors by SVM [29].

\(Y_1\) and \(Y_2\) were set as two orthonormal matrices of size \(D \times m\). The principal angles \(0 \leq \theta_1 \leq \ldots \leq \theta_m \leq \pi/2\) between the two subspaces, \(\text{span}(Y_1)\) and \(\text{span}(Y_2)\), were defined recursively by the following:

\[ \cos \theta_k = \max_{u_k \in \text{span}(Y_1)} \max_{v_k \in \text{span}(Y_2)} u_k^t v_k, \text{subject to} \]
\[ u_k^t u_k = 1, v_k^t v_k = 1, \]
\[ u_k^t v_l = 0, v_k^t v_l = 0, (i = 1, \ldots, k - 1) \]

The principal angles were computed from the SVD of \(Y_1^t Y_2\) as follows:

\[ Y_1^t Y_2 = U \cos \Theta V^t, \]

where \(U = [u_1 \ldots u_m]\), \(V = [v_1 \ldots v_m]\), and \(\cos \Theta\) is the diagonal matrix \(\cos \Theta = \text{diag}(\cos \theta_1 \ldots \cos \theta_m)\).

Based on the principal angles, the projection kernel of subspaces \(Y_1\) and \(Y_2\) was determined as follows:
\[ k_p(Y_1, Y_2) = Y_1^T Y_2^T = \text{tr}(Y_1 Y_2^T) \]

Here, we applied the projection kernel to the DMD features \( \Phi_i \) and \( \Phi_j \) to generate the Gram matrix \( k_p(\Phi_i, \Phi_j) \). To apply the Gram matrix to SVM, each component of the Gram matrix was divided by the average of all components of the Gram matrix. For comparison, the FFT features were also classified using SVM with the projection kernel. We also calculated the inner product matrix correlation \([34]\) defined as \( r_{ij}(A, B) = \frac{\text{tr}(A^T B)}{\sqrt{\text{tr}(A^T A) \text{tr}(B^T B)}} \) of two \( i \times j \) matrices \( A \) and \( B \) to examine the independent information contained in the DMD and FFT features.

2.4.3. \( T \)-distributed stochastic neighbor embedding analysis of the Gram matrix.

Based on the Grassmann kernel, the distance between each DMD mode is represented in the Grassmann space, in which the distance between each mode of the Gram matrix \( G_{i,j} \) corresponding to the inner product of \( \Phi_i \) and \( \Phi_j \) was defined as \( \text{Distance}_{i,j} = \sqrt{G_{i,i} + G_{j,j} - 2G_{i,j}} \). To visualize the distribution of each DMD mode in the space, we applied \( T \)-distributed stochastic neighbor embedding (t-SNE) to the space \([35]\).

2.5. Decoding analysis

The classifier was independently trained for each patient by SVM. The classification accuracy was evaluated by 10-fold nested cross validation. In the nested loop, the cost parameter was optimized in the range of \( 10^{-1} > \text{cost} > 10^4 \). The classification accuracies were evaluated as balanced accuracy because the number of each movement type was uneven \([36]\). The balanced accuracy is defined as the average of recall obtained in each category.

We evaluated how the classification accuracies depend on the characteristic frequency. We selected 10 modes \( \Phi_i \) \((j = 1 \ldots 10)\), which have the characteristic frequency \( \omega_i \) closest to 29 selected frequencies: 10, 20, …, 290 Hz. Then, the classification was performed with 10 selected modes for each frequency.

\( \Phi_i \) is the vector of complex values. Each complex value is expressed as a radius and a phase in a complex plane. To evaluate how much the phase information contributes to the decoding, we also evaluated the classification accuracy using the DMD features in which the phases of each \( \Phi_i \) were randomized.

2.6. Statistical tests

We compared the classification accuracies evaluated by the two types of features, DMD features and FFT features, by Student’s paired \( t \)-test, with \( p < 0.05 \) indicating significance.

We also compared the correctness of prediction by SVM trained with DMD and FFT features by Welch’s \( t \)-test, with \( p < 0.05 \) indicating significance.

The classification accuracy of each frequency was compared to the classification accuracy using the same features with shuffled labels by using Bonferroni-corrected Student’s paired \( t \)-test, with \( p < 0.05 \) indicating significance.

The phases of the DMD mode of each electrode were evaluated by the Rayleigh test to show some biases in the phase distribution of the electrode among each movement task.

3. Results

3.1. Gram matrix based on projection kernel with DMD mode and clustering of dynamics

For illustrative patient 1, we calculated the projection kernel values \( G_{i,j} \), which correspond to the inner products between each DMD mode \( \Phi_i \) and \( \Phi_j \). The projection kernel values demonstrated some clusters according to the movement types (figure 3(A)). Based on the projection kernel, we further calculated the distance matrix between each DMD mode, which represents the Grassmann space for the SVM classification. To visualize the Grassmann space, t-SNE was applied for the distance matrix to embed each trial into a two-dimensional space (figure 3(B)). Each movement type was classified into different clusters. The Grassmann space using the projection kernel seemed to be successful at classifying the movement types.

3.2. Neural decoding using DMD features

Using the Grassmann kernel, the classification of the movement types was evaluated. The classification accuracy significantly varied among the different features: DMD feature, FFT feature, DMD feature with shuffled phase, and DMD features with shuffled labels \((p = 9.12 \times 10^{-11}, \text{one-way ANOVA}; \text{figure 4})\). The three types of hand movements were successfully classified using the DMD features \((79.8 \pm 6.28\%); \text{mean} \pm 95\% \text{confidence interval})\), the FFT features \((67.3 \pm 8.56\%); \text{mean} \pm 95\% \text{confidence interval})\), and the DMD features with shuffled phase \((66.7 \pm 9.96\%); \text{mean} \pm 95\% \text{confidence interval})\), with accuracies exceeding the accuracy using the DMD features with shuffled labels \((33.1 \pm 3.90\%); \text{i.e. chance level})\). Among these classification schemes, the classification accuracy using the DMD features was significantly superior to that using the FFT features and the DMD features with shuffled phase \((p < 0.05, \text{paired Student’s} \ t\text{-test, uncorrected})\).

We also examined how the prediction of the DMD and FFT features were correlated. First, we evaluated the inner product matrix correlation of the Gram matrices of the DMD features and the FFT features. For 10 out of 11 patients, the correlation was more than 0.8 (supplementary information, table 2), suggesting that the Gram matrices were similar to each other. However, for patient 5, whose correlation was 0.359, the classification accuracy was higher with the DMD features for this patient. Second, we compared the predictions using two features. We divided the classification results with the two features into three
Figure 3. Gram matrix and similarity in the Grassmann space for patient 1. (A) The projection kernel values were color-coded on the matrix of each trial. The numbers on the vertical and horizontal axes represent the movement type (1, grasping; 2, opening; 3, pinching). (B) The distribution of the data for each trial in the Grassmann space that was embedded into two dimensions by t-SNE. Each class corresponds to the movement type.

Figure 4. Classification accuracy using different features. The mean and 95% confidence interval of the classification accuracies are shown for the classification using the DMD features, FFT features, DMD features with shuffled phase, and DMD features with shuffled labels (*p < 0.05, **p < 0.01, paired Student’s t-test).

groups for each patient: Group 1, the predictions of both features were the same; Group 2, the prediction of DMD was correct, but that of FFT was incorrect; and Group 3, the prediction of DMD was incorrect, but that of FFT was correct. Among all patients, the ratio of the groups was significantly high for Group 1 (66.5 ± 8.02%) (figure 5). Moreover, for the conflicting predictions, Group 2 was significantly larger than Group 3. Therefore, these two features mostly result in the same predictions, but in the case of conflicting predictions, the DMD features predicted significantly better than the FFT features.

We compared the classification accuracy using subsets of the DMD features to evaluate which frequency band modes contributed to the classification. The classification accuracies using DMD features of the characteristic frequency close to the selected frequency significantly exceeded the accuracy by chance at 60, 80, 100, 110, and 150 Hz (p < 0.01, Bonferroni-corrected paired Student’s t-test; figure 6). The peak mean accuracy was 50.4 ± 2.42% (mean ± 95% confidence interval) at 100 Hz, which corresponds to the high gamma band.

3.3. DMD modes during movement

The DMD mode captured spatiotemporal patterns characterizing movement type. For the high gamma band frequency (80–150 Hz) in patient 1, the averaged amplitudes and phases of the DMD mode were evaluated and color-coded according to electrode position (figure 7(A)). The magnitude of the DMD mode was high at the electrodes on the primary motor cortex and demonstrated different patterns in accordance with movement type for patient 1. In addition, the relative phases of each electrode were significantly biased around the motor cortex, and the averaged relative phases were different among movement types. Similarly, the magnitudes and phases of the DMD modes changed based on the movement types for all other patients (supplementary information).
Figure 5. Comparison of correctness of prediction by SVM trained with DMD and FFT features. The mean and 95% confidence interval of prediction ratio where the two sets of features had the same prediction label, where DMD had correct prediction labels and FFT had incorrect prediction labels, and where DMD had incorrect prediction labels and FFT had correct prediction labels (**p < 0.01, Bonferroni-corrected Welch’s t-test).

Figure 6. Classification accuracy for each frequency mode. The classification accuracies using the DMD features whose frequency components were close to each frequency point (blue line) were compared to those using the same features with shuffled labels (chance level, orange line). The accuracies at approximately 100 Hz were significantly higher than those at chance level (*p < 0.05, **p < 0.01, paired Student’s t-test, Bonferroni corrected).

On the other hand, although the FFT powers of each electrode averaged among the same frequency band demonstrated some spatial pattern characterizing the movement type, the distribution of the high powers was different than that of the magnitude of the DMD mode and not restricted around the primary motor cortex (figure 7(B)). The DMD mode captured some characteristic spatiotemporal patterns of the phases and magnitudes.

4. Discussion

Our analyses indicated that the DMD mode with projection kernel provided better accuracy of the neural decoding than the FFT power, thereby demonstrating that DMD succeeded in capturing neural information better than FFT. Moreover, the projection kernel was suggested to represent the characteristic differences in ECoG signals among different movement types. Our
Figure 7. DMD features and FFT features for patient 1. (A) The top row represents the averaged magnitudes of the DMD modes at each electrode. These magnitudes were obtained by averaging the mode magnitudes in the high gamma band among each movement type. The bottom row represents the averaged phases of the DMD modes. The phase was evaluated for the modes with positive imagery components. Notably, each DMD mode was composed of a pair of complex conjugates. The phase of each electrode was subtracted by the phase of electrode number 1 (upper right of 60 displayed electrodes). The relative phases were averaged among each movement type. The relative phases at the electrodes with black circles were significantly biased ($p < 0.05$, uncorrected Rayleigh test). The white dotted line represents the location of the central sulcus. (B) The scaled FFT power was averaged between 80 and 150 Hz (high gamma band) of each movement type and color coded by electrode position.

proposed method using DMD with a projection kernel is a promising technique for neural decoding.

DMD succeeded in extracting some spatial and temporal patterns characterizing three types of hand movements. Although a previous study demonstrated that DMD extracts some characteristic spatial patterns in ECoG signals [27], it remained unclear whether these patterns included more information about the movement types than the patterns evaluated by FFT. In our study, the decoding accuracy using DMD features was higher than that using FFT and DMD features with shuffled phases. Notably, both DMD and FFT were able to capture spatial and temporal patterns of each frequency range, as illustrated in figure 4. However, FFT was not able to maintain the information of phases between the dynamics of different channels. Indeed, the classification accuracy using the DMD mode significantly decreased when the phases of the DMD modes were shuffled and were similar to the accuracy using the FFT. Moreover, comparing the predictions of the DMD features and the FFT features, these predictions mostly overlapped, but the DMD features had more useful information than the FFT features. Therefore, our data suggest that the DMD mode improved the classification accuracy compared to the FFT by capturing some spatially and temporally coherent patterns in the ECoG signals to obtain motor information.

Among the frequency features examined, the DMD modes at approximately 100 Hz demonstrated higher classification accuracy, implying that they contained a larger amount of motor information. Some studies using FFT powers have also shown that the frequency powers at approximately 100 Hz provide more motor information [15, 37]. Here, DMD was suggested to extract the spatial and temporal pattern relating to the oscillation at 100 Hz. Moreover, our results demonstrated that the DMD mode close to 60 Hz was also decoded with accuracy exceeding the chance level. Although a previous study using FFT signals demonstrated that the powers at approximately 60 Hz were difficult to decode due to some electrical power noise [37], our results suggested that the DMD mode is useful for extracting neural information from the ECoG signals, even with noise.

There are some problems in applying the proposed decoding method to a BCI. First, the DMD mode calculation must be made within at least several hundred milliseconds, which would require sufficient machine power and an appropriate approximation of the calculation. An incremental SVD might
be useful to obtain such an appropriate approximation [38–40]. Second, we might need to select an appropriate DMD mode for online use during which the activity patterns of the brain continuously change. Because DMD assumes a stationary nonlinear property over a certain period of time, even for non-stationary dynamics, adapting the method to non-stationary data might be difficult. Therefore, some other machine learning techniques, such as a hidden Markov model, may be needed to address non-stationary states [41].

Although some technical improvements (e.g. calculation speed, adapting to nonstationary data) are necessary, improvements in the decoding accuracy of the performed movement types will improve the BCI as a communication device for the patient with limited ability to move their bodies. Our previous study demonstrated that the ECoG signals between the performed movement and attempted movements of paralyzed patients were similar and were decoded similarly. Our proposed method is also expected to improve the performance of attempted movements. Therefore, our proposed method with DMD will improve the decoding accuracy for paralyzed patients. Using precise decoding of attempted movements, paralyzed patients will be able to control some external devices to communicate with others [10]. Our method will improve BCI performance to realize better communication using BCI for paralyzed patients.

Practically, our method can be applied for to the functional mapping of movements or other various cortical functions, such as language, which has been developed with power spectral analysis of ECoG signals [42]. DMD will reveal the spatially and temporally coherent pattern of the ECoG signals in the characteristic frequency range with the highest information regarding particular cortical functions and might improve the accuracy of the functional mapping compared to that using power spectral analysis without phase coherence among electrodes. Similarly, our method can be applied to identify particular characteristic ECoG signals related to epileptic activity. Our neural decoding with DMD will be useful to identify spatially and temporally coherent patterns that represent particular types of information.

This study had some limitations. To simplify and shorten the calculation, we aligned the mode magnitude by the L2 norm instead of the orthogonalization. Therefore, the projection kernels between two matrices were not guaranteed to be the same when the rows of the matrix were exchanged with each other [29]. However, our results demonstrated that the simplified method succeeded in decoding three movement types with accuracy that exceeded that using FFT. The proposed method with a projection kernel was shown to be useful for motor decoding. Our future work should evaluate how orthogonalization affects the decoding accuracy. In addition, our current method ignored the frequency information. Although we used the DMD modes at specific frequencies, we did not use the frequency components for the decoding. The decoding accuracy is expected to improve with the addition of the frequency components into the DMD features. Notably, we compared the classification accuracy of three types of movements that were different among patients. The differences in the movement types might have affected the accuracy. However, by comparing the accuracy using FFT features and the DMD features with Grassmann kernels for each patient, we eliminated the effects of the different sets of movement types. Moreover, the ECoG signals in this study came from two types of patients with and without paralysis in their hands. These patients had different diseases and were implanted with different numbers of electrodes but implantations on their sensorimotor cortices was common among the patients. Notably, those differences induced no apparent differences in the decoding accuracies (supplementary information). However, the anatomical pattern of the DMD features might be affected by ischemic stroke and brain tumors in the motor cortex. We have demonstrated that the classification accuracy was significantly improved by DMD features compared to FFT features using the same classification method for the ECoG signals from the sensorimotor cortex in various conditions.

5. Conclusions

We demonstrated that DMD, which extracts spatial and temporal patterns with phase relations, is useful for extracting more motor information than FFT. The proposed method will be applied to improve BCIs and to capture some characteristic patterns of cortical activities related to motor functions.

Study funding

This research was conducted under the Japan Science and Technology Agency (JST) Core Research for Evolutional Science and Technology (JPMJCR18A5). This research was also supported in part by the Japan Science and Technology Agency (JST) Pre-cursory Research for Embryonic Science and Technology (JPMJPR1506), Exploratory Research for Advanced Technology (JPMJER1801) and CREST (JPMJCR1913); Grants-in-Aid for Scientific Research from KAKENHI (JP17H06032, JP15H05710, JP18H05522 and JP18H03287); Grants from the Japan Agency for Medical Research and Development (AMED) (19dm0207070h0001, JP19dm0307103, JP19de0107001, JP18H05522 and JP19dm0307009); Contract research with the National Institute of Information and Communications Technology (Grant #209); SIP (Innovative AI Hospital System) of NIBIOHN; and the Canon Foundation.
Conflict of interest

We declare no competing interests.

Author contributions

Yoshiyuki Shiraishi: Software, Validation, Formal analysis, Investigation, Visualization, Writing-Original Draft, Writing-Review and Editing. Yoshihiro Kawahara: Conceptualization, Methodology, Writing-Review and Editing. Okito Yamashita: Conceptualization, Writing-Review and Editing. Ryohei Fukuma: Software, Writing-Review and Editing. Shota Yamamoto: Data Curation. Youshi Saihoh: Resources. Haruhiko Kishima: Resources. Takafumi Yanagisawa: Conceptualization, Investigation, Data Curation, Writing-Original Draft, Writing-Review and Editing. Supervision, Project administration, Funding acquisition.

Data/code availability

The data that support the findings of this study are not publicly available because they contain information that could compromise research participants’ privacy and/or consent. The code used in this study are publicly available at https://github.com/yanagisawa-lab.

ORCID ID

Takafumi Yanagisawa  https://orcid.org/0000-0002-2057-0612

References

[1] Wolpaw J R, Birbaumer N, McFarland D J, Pfurtscheller G and Vaughan T M 2002 Brain–computer interfaces for communication and control Clin. Neurophysiol. 113 767–91
[2] Kageyama Y et al 2014 Severely affected ALS patients have broad and high expectations for brain-machine interfaces Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration 15 513–9
[3] Leuthardt E C, Schalk G, Wolpaw J R, Ojemann J G and Moran D W 2004 A brain–computer interface using electrocorticographic signals in humans J. Neural Eng. 1 63–71
[4] Hochberg L R et al 2012 Reach and grasp by people with tetraplegia using a neurally controlled robotic arm Nature 485 372–5
[5] Pandarinath C, Gilja V, Blabe C H, Nuyujukian P, Sarma A A, Sorice B L, Eskandar E N, Hochberg L R, Henderson J M and Shenoy K V 2015 Neural population dynamics in human motor cortex during movements in people with ALS Elife 4 e07436
[6] Anumanchipalli G K, Chartier J and Chang E F 2019 Speech synthesis from neural decoding of spoken sentences Nature 568 493–8
[7] Chao Z C, Nagasaka Y and Fujiw N 2010 Long-term asynchronous decoding of arm motion using electrocorticographic signals in monkeys Front. Neuroeng. 3 3
[8] Fernandez E, Gregor B, House P A, Aranda I, Botella C, Albissu J, Soto-sánchez C, Alfaro A and Normann R A 2014 Acute human brain responses to intracortical microelectrode arrays: challenges and future prospects Front. Neuroeng. 7 24
[9] Morris S, Hirata M, Sugata H, Goto T, Matsuishi K, Yanagisawa T, Saiho Y, Kishima H and Yoshimine T 2015 Patient-specific cortical electrodes for sulcal and gyral implantation IEEE Trans. Biomed. Eng. 62 1034–41
[10] Vansteenberg M J et al 2016 Fully implanted brain–computer interface in a locked-in patient with ALS New Engl. J. Med. 375 2090–6
[11] Pandarinath C, Nuyujukian P, Blabe C H, Sorice B L, Saab J, Willett F R, Hochberg J R, Shenoy K V and Henderson J M 2017 High performance communication by people with paralysis using an intracortical brain–computer interface Elife 6
[12] Miller K J, Zanos S, Fetz E E, den Nijs M and Ojemann J G 2009 Decoupling the cortical power spectrum reveals real-time representation of individual finger movements in humans J. Neurosci. 29 3132–7
[13] Flamary R and Rakotomamonjy A 2012 Decoding finger movements from ECoG signals using switching linear models Front. Neurosci. 6 29
[14] Freedenburg Z V et al 2019 Sensormotor ECoG signal features for BCI control: a comparison between people with locked-in syndrome and able-bodied controls Front. Neurosci. 13 1038
[15] Yanagisawa T, Hirata M, Saiho Y, Goto T, Kishima H, Fukuma R, Yokoi H, Kamitani Y and Yoshimine T 2011 Real-time control of a prosthetic hand using human electrocorticography signals J. Neurosurg. 114 1715–22
[16] Yanagisawa T, Hirata M, Saiho Y, Kato A, Shibuya D, Kamitani Y and Yoshimine T 2009 Neural decoding using gyral and intrasulcal electrocorticograms Neuroimage 45 1099–106
[17] Yanagisawa T, Hirata M, Saiho Y, Kishima H, Matsuishi K, Goto T, Fukuma R, Yokoi H, Kamitani Y and Yoshimine T 2012 Electro cortical control of a prosthetic arm in paralyzed patients Ann. Neurol. 71 353–61
[18] Miller K J 2019 A library of human electrocorticographic data and analyses Nat. Hum. Behav. 3 1225–35
[19] Wang W et al 2013 An electrocorticographic brain interface in an individual with tetraplegia PLoS One 8 e55344
[20] Volkova S, Lebedev M A, Kaplan A and Ossaditch A 2019 Decoding movement from electrocorticographic activity: a review Front. Neuroinform. 13 74
[21] Rowley C W, Mezic I, Bagheri S, Schlatter P and Henningson D S 2009 Spectral analysis of nonlinear flows J. Fluid Mech. 641 115–27
[22] Schmid P 2010 Dynamic mode decomposition of numerical and experimental data J. Fluid Mech. 656 5–28
[23] Mezić I 2013 Analysis of fluid flows via spectral properties of the Koopman operator Ann. Rev. Fluid Mech. 45 357–78
[24] Sarkar S, Ganguly S, Dalal A, Saha P and Chakraborty S 2013 Mixed convective flow stability of nanofluids past a square cylinder by dynamic mode decomposition Int. J. Heat Fluid Flow 44 624–34
[25] Casorso J, Kong X, Chi W, Van De Ville D, Yeo B T T and Liegeois R 2019 Dynamic mode decomposition of resting-state and task fMRI Neuroimage 194 42–54
[26] Berger E, Sastuba M, Vogt D, Jung B and Ben Amor H 2015 Estimation of perturbations in robotic behavior using dynamic mode decomposition Adv. Robot. 29 331–43
[27] Brunton B W, Johnson L A, Ojemann J G and Kutz J N 2016 Extracting spatial-temporal coherent patterns in large-scale neural recordings using dynamic mode decomposition J. Neurosci. Methods 258 1–15
[28] Kutz J N, Fu X and Brunton S L 2016 Multiresolution dynamic mode decomposition SIAM J. Appl. Dyn. Syst. 15 713–35
[29] Hann M and Lee D D 2008 Grassmann discriminant analysis: a unifying view on subspace-based learning Proc. 25th Int. Conf. on Machine Learning pp 376–83
[30] Kubanek J, Miller K J, Ojemann J G, Wolpaw J R and Schalk G 2009 Decoding flexion of individual fingers using...
electrocorticographic signals in humans J. Neural Eng. 6 066001

[31] Reddy C G, Reddy G G, Kawasaki H, Oya H, Miller L E and Howard M A 2009 Decoding movement-related cortical potentials from electrocorticography Neurourg. Focus 27 E11

[32] Gunduz A, Brunner P, Daitch A, Leuthardt E C, Ritaccio A L, Pesaran B and Schalk G 2012 Decoding covert spatial attention using electrocorticographic (ECoG) signals in humans Neuroimage 60 2285–93

[33] Thomas T M, Candrea D N, Fifer M S, McMullen D P, Anderson W S, Thakor N V and Crone N E 2019 Decoding native cortical representations for flexion and extension at upper limb joints using electrocorticography IEEE Trans. Neural Syst. Rehabil. Eng. 27 293–303

[34] Ramsay J O, ten Berge J M F and Stryan G P H 1984 Matrix correlation Psychometrika 49 403–23

[35] van der Maaten L and Hinton G 2008 Visualizing data using t-SNE J. Mach. Learn. Res. 9 2579–605

[36] Brodersen K H, et al 2010 The balanced accuracy and its posterior distribution. Int. Conf. Pattern Recognition pp 3121–4

[37] Pistohl T, Schulze-Bonhage A, Aertsén A, Mehring C and Ball T 2012 Decoding natural grasp types from human ECoG Neuroimage 59 248–60

[38] Brand M 2002 Incremental singular value decomposition of uncertain data with missing values. Computer Vision - ECCV 2002 2350 707–20

[39] Ghashami M, Liberty E, Phillips J M and Woodruff D P 2016 Frequent directions: simple and deterministic matrix sketching Siam J. Comput. 45 1762–92

[40] Zhou X, He J, Huang G and Zhang Y 2015 SVD-based incremental approaches for recommender systems J. Comput. Syst. Sci. 81 717–33

[41] Wissel T, Pfeiffer T, Frohlich R, Knight R T, Chang E F, Hinrichs H, Rieger J W and Rose G 2013 Hidden Markov model and support vector machine based decoding of finger movements using electrocorticography J. Neural Eng. 10 056020

[42] Crone N E, Sinai A and Korzeniewska A 2006 High-frequency gamma oscillations and human brain mapping with electrocorticography Prog. Brain Res. 159 275–95