Research on Feature Extraction Algorithm Commonly Used in Brain-computer Interface Technology

Yifan Zhang¹*, Yaojun Wang²

¹Electronic information and technology, Miami college of Henan University, Kaifeng, Henan, 475000, China
²Electronic Science and technology, Physics Institute of Zhengzhou University, zhengzhou, Henan, 450000, China

*Corresponding author’s e-mail: 171840776@snnj.edu.cn

Abstract. Brain-computer interface (BCI) is an effective and direct channel of information exchange between human brain and external devices such as computer, which can provide auxiliary information acquisition and treatment means for the medical and other fields in the future. This paper focuses on four kinds of feature extraction algorithms such as power spectral density (PSD), wavelet transform, Hilbert-Huang transform (HHT) and common space pattern (CSP), which are commonly used to process abnormal electroencephalogram (EEG) signals in brain-computer interface technology. This paper also introduces their respective principles, characteristics and application fields, analyzes and compares the advantages and disadvantages of these algorithms, and obtains the development direction of feature extraction algorithms in the future. Finally, it also briefly discusses the ethical issues brought about by brain-computer interface technology.

1. Introduction

The brain-computer interface consists of two main parts: brain and machine. The brain is the nervous system of a person or an animal that controls our activities. Machine refers to the equipment collecting signals, such as specific algorithms, data processing, image generation and other external devices. Theoretically, through brain-computer interface, we can realize the non-"language" communication between human and human, human and animal, that is, information transmission can happen through machinery equipment.

The brain-computer interface can be divided into invasive brain-computer interface, partial invasive brain-computer interface and non-invasive brain-computer interface according to the different ways of collecting signals. The main purpose of the intrusive brain-computer interface is to realize the reconstruction of people's special senses and the recovery of motor function. It can collect better quality signals, but it also has problems difficult to solve, such as short life span and the possibility to produce immune response to human health. Partial the invasive brain-computer interface is to implant the signal-collecting electrodes into the cranial cavity, which has fewer problems than the invasive type. Non-invasive brain-computer interface is least harmful to human body, but it also brings problems such as more noise signals and difficulty in post-processing data.

As a result, a variety of EEG systems (BCI) have been designed. BCI refers to the combination of functions for the collection, processing and classification of EEG in most cases. First, EEG signals need to be collected with external information collection equipment to get the initial input signals.
Then we need to process the information and convert the analog signal into a digital signal with parameter characteristics, so that the computer can analyze it. Through the analysis, we can get the action command and thought activities of the observed object, and then realize the control of the external object, information exchange and so on.

2. Study on Feature Extraction Algorithm

Signal processing and pattern recognition algorithm are the key techniques in BCI system. The processing of EEG signal includes three steps: pre-processing, feature extraction and pattern classification. By pre-processing all kinds of noise interference signals in EEG are eliminated, and the EEG signals related to thinking tasks are strengthened, then feature pattern is extracted with effective feature extraction algorithm, and finally classified and identified with pattern classifier. This paper mainly discusses the feature extraction algorithm applicable to EEG signal.

Fourier transform is a classical signal analysis method, but it is based on the idea of uncertainty principle. There is an inevitable contradiction, that is, accurate information both in time domain and frequency domain cannot be obtained at the same time. In other words, Fourier analysis is only suitable for frequency-stable signals, not for non-stationary signals. At the same time, as a global change, Fourier transform does not have the ability of localization analysis. Due to the non-linear and non-stationary characteristics of EEG, Fourier analysis is no longer applicable, and people are constantly seeking new analytical methods to process EEG.

Some commonly used feature extraction algorithms are analyzed and compared now.

2.1. Power spectral density principle

Power spectral density analysis is a common frequency domain analysis method. EEG signal is a kind of non-stationary random signal. Generally, the duration of random signal is infinite, so its total energy is infinite and cannot be processed by traditional Fourier transform. Although the total energy of the random signal is infinite, the average power of the random signal is limited, so it is meaningful to analyze the frequency domain of the random signal from the power spectrum. As a result, power spectral density (PSD) is often used to analyze the frequency domain characteristics of EEG signals.

The basic theory of power spectral density is described as below:

Power spectral density \( S(f) \) is a mapping with frequency \( f \) as independent variable, which reflects how much power a signal has in the frequency component \( f \).

We presume a random process \( X(t) \) and define a truncation threshold \( t_0 \), and then the truncation process \( X \) of the stochastic process \( X_{t_0}(t) \) can be defined as:

\[
    X_{t_0}(t) = \begin{cases} 
    X(t) & |t| \leq t_0 \\
    0 & |t| > t_0 
    \end{cases}
\]

The energy of the stochastic process can be defined as:

\[
    E_{X_{t_0}} = \int_{-t_0}^{t_0} X^2(t)dt = \int_{-t_0}^{t_0} X^2_{t_0}(t)dt
\]

The average power can be obtained with the derivation of energy function:

\[
    P_{X_{t_0}} = \frac{1}{2t_0} \int_{-t_0}^{t_0} X_{t_0}^2(t)dt
\]

According to Parseval theorem (energy is equal in both time and frequency domains) we can obtain:

\[
    P_{X_{t_0}} = \frac{1}{2t_0} \int_{-t_0}^{t_0} |X_{t_0}(f)|^2 df
\]

Here \( X_{t_0}(f) \) is the Fourier transformed formed of \( X_{t_0}(t) \). Since the random process \( X \) is limited in a finite time interval \([-t_0,t_0]\), its Fourier transform will no longer be limited. In addition, because
The average power $P_{X_0}$ is a random variable, to get the final overall average power, it is necessary to get the expected value of the random variable:

$$\overline{P_X} = \lim_{t_0 \to \infty} \frac{1}{2t_0} \int_{-\infty}^{\infty} E[|X_0(f)|]df = \int_{-\infty}^{\infty} \lim_{t_0 \to \infty} \frac{E[|X_0(f)|]}{2t_0} df$$

The integral function in the expression is extracted separately and is defined as $S(f)$:

$$S(f) = \lim_{t_0 \to \infty} \frac{E[|X_0(f)|]}{2t_0}$$

The average power $\overline{P_X}$ can be expressed as $\int_{-\infty}^{\infty} S(f)df$. It can be seen that the function $S(f)$ can represent the power of the frequency component of each minimum limit unit, so $S(f)$ will be called the power spectral density.

The power spectrum analysis reflects the frequency component of the signal and the relative degree of each component. The significance of analyzing EEG signal with power spectrum is that the amplitude signal is transformed into the spectrum of the relative change of success rate and frequency through the amplitude energy conversion of signal with time. The spectrum contains a lot of useful information, through which we can intuitively obtain the distribution and variation of various rhythms in EEG signals. The powerful computing ability of computer and the appearance of discrete Fourier transform (DFT) FFT make power spectrum analysis widely used.

But the power spectrum analysis also has certain limitations like the possibility of losing all phase information, so it is impossible to recover the original signal through the power spectrum. The recognition ability of power spectral density will also change when the signal-to-noise ratio changes. Because the basic features extracted from time domain and frequency domain are sensitive to the change of signal-to-noise ratio, it is difficult to improve the classification recognition ability of this method when the signal-to-noise ratio is unknown.

2.2. Wavelet Transform Principle

Frequency domain analysis method will cause phase loss and it is unable to recover the original signal directly. In order to solve this problem, the time-frequency analysis method represented by wavelet transform becomes the mainstream gradually. In 1910, the idea of orthogonal basis of wavelet norms was first proposed by Harr, and the concept of "wavelet analysis" was first proposed by Morlet in 1981, so it is called Morlet wavelet. In 1984, Morlet and Grossman jointly proposed the mathematical model of wavelet analysis. Wavelet transform is developed on the basis of STST. It inherits the idea of short-time Fourier transform localization. It has better localization characteristics in time domain and frequency domain has the ability of multi-resolution analysis at the same time. It overcomes the shortcoming of window size not changing with frequency. It can provide a "time-frequency" window which changes with frequency and has good self-adaptive ability. It is an ideal tool for signal time-frequency analysis and processing, especially for image and signal denoising.

The basic principles of wavelet transform are as follows:

① Unlike Fourier coefficients, wavelet coefficients vary with frequency, and for the same frequency index $j$, the wavelet coefficients of $k$ are also different at different times.

② Lower time resolution and higher frequency resolution are used in the low frequency part of the signal (the signal is stable), and lower frequency resolution and higher time resolution are used in the high frequency part (the frequency transformation is not big).

③ The width of the time window obtained through wavelet transform is inversely proportional to the frequency of the signal, that is, the time window becomes narrower when high frequency signal is detected, and the time window widens when low frequency signal is detected, which realizes the change of the window width with frequency.
In terms of basis functions, different from Fourier analysis, the basis functions of wavelet analysis are not unique and can be chosen flexibly and constructed according to the problems to be solved.

For arbitrary function \( f(t) \in L^1(\mathbb{R}) \), the formula of continuous wavelet transform is as follow:

\[
W_f(a,b) = \langle f, \psi_{a,b} \rangle = \left| a \right|^{-1/2} \int f(t) \psi\left(\frac{t-b}{a}\right)dt
\]

Where, \( \psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right), a, b \in \mathbb{R}, a \neq 0 \) is called a wavelet sequence, \( a \) is the scaling factor and \( b \) is the smoothing factor. The inverse transformation formula is as follow:

\[
f(t) = \frac{1}{C} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_f(a,b) \psi\left(\frac{t-b}{a}\right) a \ da \ db
\]

In practical application, continuous wavelet must be discretized, and the coefficient of discrete wavelet transform is as follow:

\[
f(t) = \sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} C_{j,k} \psi_{j,k}(t)
\]

The discrete wavelet function is:

\[
\psi_{j,k}(t) = a_0^{-j/2} \psi\left(\frac{t-kb_0}{a_0^j}\right) = a_0^{-j/2} \psi(a_0^{-j}t - kb_0)
\]

The inverse transformation formula is:

\[
f(t) = C \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} C_{j,k} \psi_{j,k}(t)
\]

In the formula, \( C \) is a signal-independent constant.

Wavelet transform has the following advantages in the application of image processing:

1. Wavelet decomposition can cover the entire frequency domain, which provides a complete mathematical description.
2. Wavelet transform can greatly reduce or remove the correlation among different features by selecting suitable filters.
3. Wavelet transform has the characteristic of "zoom". In low frequency section, high frequency resolution and low time resolution can be used (wide analysis window), and in high frequency section, low frequency resolution and high time resolution can be used (narrow analysis window).
4. There is a fast algorithm - Mallat wavelet decomposition algorithm for wavelet transform.

Wavelet analysis has the advantages of stable bandwidth, multi-scale and so on. Selecting suitable basic wavelet can simultaneously reflect some characteristics of signal simultaneously in time-frequency domain. The essence of wavelet transform is a band-pass filter, whose central frequency is determined by the size of scale transformation, so it can detect the abnormal wave components in EEG signal effectively.

2.3. Principle of Hilbert-Huang transform

Hilbert-Huang transform is also a common time-frequency analysis method, which can process EEG signal effectively. In 1998, the empirical mode method proposed by Norden E. Huang, a Chinese-origin scientist, can be used to process the instantaneous frequency, based on which Hilbert-Huang transform is proposed.

In general, the single-valued functions in the collected signals are time-dependent and contain a variety of oscillatory modes, so it is impossible to carry out Hilbert transform directly. Therefore, before Hilbert transform it is necessary to separate these internal modes and get a function with only one mode, i.e., the intrinsic mode function (IMF). The method is to use empirical mode decomposition (EMD) to separate the original data into the accumulation of wave sequences with different
characteristics and select the IMF branches. Huang et al. offer two conditions that the IMF must meet: narrow bandwidth and symmetry. Narrow bandwidth shows that the difference between the pole and the zero number is less than or equal to 1 in all data ranges, and symmetry shows that at any time point, the mean value of the maximum and the minimum of the function is zero. The available instantaneous parameters can be obtained after the IMF which meets the conditions is Hilbert transformed.

The basic methods of Hilbert transform are as follows:

Set $X(t)$ as any time series signal, and $Y(t)$ is the Hilbert transform of this signal, then $Y(t)$ can be expressed by $X(t)$ as:

$$Y(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{X(\tau)}{t - \tau} d\tau$$

At the same time, $X(t)$ can be expressed by $Y(t)$ as:

$$X(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{Y(\tau)}{t - \tau} d\tau$$

As it can be seen from the above formulas, $X(t)$ and $Y(t)$ form complex conjugate pairs, which are all related to time, showing certain signal features. And new analytical signals can be obtained by various mathematical transformations:

$$Z(t) = X(t) + jY(t) = A(t)e^{j\theta(t)}$$

As it can be seen from the above formula, the three most important instantaneous feature parameters in Hilbert transform are instantaneous amplitude, instantaneous phase and instantaneous frequency. Their common point is that they can be converted to each other through certain mathematical expressions:

$$A(t) = \sqrt{X^2(t) + Y^2(t)}$$

$$\theta(t) = \arctan \left( \frac{Y(t)}{X(t)} \right)$$

$$f(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt}$$

Therefore, Hilbert-Huang transform can truly extract the instantaneous parameters, which is very important for the analysis of nonlinear and non-stationary signals. As a new method of signal time-frequency localization, Hilbert-Huang transform has many advantages:

① Hilbert-Huang transform is easy to use and has good self-adaptive ability. No basis function is required and the empirical mode decomposition can be performed according to the characteristics of the signal itself.

② Hilbert-Huang transform has a good time-frequency concentration, avoids uncertainty principle, can show the accurate relationship between time and frequency, and improves the resolution of time and frequency.

However, there are also some problems that cannot be ignored in this algorithm. For example, when sampling at the high-frequency part of the continuous signal, there is because the sampling rate fails to meet the Nyquist sampling theorem, which is called endpoint effect; the IMF with discontinuous signals contains very different characteristic time scales, that is, modal aliasing effect, which does not meet the demand in this case; and the phenomenon of negative frequency which has no physical meaning will occurs when calculating the instantaneous frequency.

2.4. Principle of common space pattern algorithm

The common space pattern (CSP) was first proposed by Fukunage et al., and was later used in BCI by Romeser and his colleagues. CSP is a kind of spatial filtering feature extraction algorithm under two classification tasks. Because of the difference of different signals in energy and frequency, the projection on a spatial filter can be obviously different, and each kind of spatial distribution component can be extracted from the multi-channel brain-computer interface data. The basic principle is to use the diagonalization of matrix to find a set of optimal spatial filters for projection, so that the difference between the two kinds of signals can be maximized, and the feature vectors with high discrimination can be obtained.
The specific process of CSP algorithm is as follows:

1. Segment the raw data by category. Sample data $E$ can be divided into $E_1$ and $E_2$.

2. Covariance matrix calculating the original data after segmentation:

$$C_i = \frac{E_i \cdot E_i^T}{\text{trace}(E_i \cdot E_i^T)} (i = 1, 2)$$

$\text{trace}(E)$ indicates the trace of the matrix $E$.

Respectively calculate the covariance matrices of the original data, and $C_1$ and $C_2$ are the expectation of the spatial covariance matrix of the first and second kinds of sample data respectively. $C_c$ represents the sum of the spatial covariance matrix of the two kinds of data, and then we can get $C_c = C_1 + C_2$.

3. Conduct orthogonal whitening transformation and diagonalization at the same time. $C_c$ is a positive definite matrix, derived from the singular decomposition theorem:

$$C_c = U_c \Lambda_c U_c^T$$

$U_c$ is a matrix of eigenvectors, $\Lambda_c$ is a diagonal matrix of eigenvalues, and the eigenvalues are arranged in descending order. With the albino conversion of $U_c$, we can obtain:

$$P = \frac{1}{\sqrt{\Lambda_c}}, U_c^T$$

By applying matrix $P$ to $C_1$ and $C_2$, we can obtain:

$$S_1 = P C_1 P^T, S_2 = P C_2 P^T$$

$S_1$ and $S_2$ have a common eigenvector, and there are two diagonal matrices $\Lambda_1$ and $\Lambda_2$, and an eigenvector matrix $B$, which satisfy the following conditions:

$$S_1 = B \Lambda_1 B', S_2 = B \Lambda_2 B', \Lambda_1 + \Lambda_2 = I$$

Where, $I$ is a unit matrix. Thus, the sum of the eigenvalues of $S_1$ and $S_2$ is $I$.

4. Calculate the projection matrix. For the eigenvector matrix $Q$, when one category $S_1$ has the maximum eigenvalue, the other category $S_2$ has the minimum eigenvalue at the same time, so we can use matrix $Q$ to realize the classification of two kinds of problems, and obtain the projection matrix:

$$W = (Q^T P)^T$$

5. The eigenvector matrix is obtained by projection. The original EEG data $E_{(M \times N)}$ are projected through the projection matrix $W$ to obtain the feature matrix:

$$Z_{M \times N} = W_{M \times N} E_{M \times N}$$

The first $m$ lines and the last $m$ lines of $Z_{M \times N} (2m < M)$ can be selected as the eigenvalues of the original input data.

6. Normalization of eigenvalues.

$$y_i = \log \left( \frac{\text{var}(Z_i)}{\sum_{i=1}^{m} \text{var}(Z_i)} \right)$$

Where, $y_i$ is the normalized eigenvector matrix of the $i$ sample.

The information of CSP eigenvector matrix generated by the algorithm is not equivalent. The feature information is mainly concentrated in the head and tail of the feature matrix, but the middle feature information can be ignored, so the data from the first $m$ lines and the last $m$ lines are selected as the eigenvector matrix.

Spatial filter is very suitable for processing multi-dimensional EEG signals and data. It can use the spatial correlation of EEG signals synchronously, eliminate the signal noise and locate the nerve
activity of local cortex. By combining spatial filtering with time domain and frequency domain features, better processing results can be obtained.

Because the information of different features is contained in different frequency sections of the signal, EEG can be separated into different specific frequency sections for analysis, which maximizes the variance of the inter-class data and minimizes the variance of the intra-class data. It can be said that CSP is one of the best and most widely used feature extraction algorithms. But there are also some problems in the common-space pattern algorithm, such as over-fitting phenomenon caused by many EEG channels and insufficient training data and frequency band selection.

3. Ethical issues
In the process of studying the algorithm, we found a series of ethical issues brought by brain-computer interface. How to settle these ethical issues will be another challenge for the development of brain-computer interface in the future.

All the above feature extraction algorithms require clear and low noise signals, but the EEG signals collected by non-invasive devices are usually weak and inconvenient to process. In order to obtain high-quality EEG signals, invasive devices are sometimes used to implant electrodes into the intracavitary cortex, which means greater trauma and risk. For example, local mechanical damage to brain tissue may occur during electrode implantation; rejection may occur after surgery, resulting in brain damage; there is no reliable data on the use time after the electrode implantation. There is no clear answer to how such trauma can be repaired, who should take the risk, and whether it is reasonable to add insurance to the cost of the purchase.

The privacy issues associated with brain-computer interfaces should not be overlooked. Brain electrical signals carry rich personal information, and these signals are collected and analyzed, and people's behaviors are interfered with, which will inevitably infringe upon personal privacy. If a large amount of personal information falls into the hands of lawbreakers, it may be illegally disseminated and traded, or even used to intervene and control the user's consciousness and behavior, causing great harm to the user's personal interests and social and public interests. In addition, whether the brain-computer interface has the right to interfere with the free will of human beings, what is the standard of interference, that makes it, and whether it conforms to ethics and laws and regulations, all need to be considered by the whole human society for a long time.

4. Conclusion
The key of brain-computer interface technology is the processing of EEG signal, and feature extraction is an important step in EEG processing. After comparison, it is found that several common feature extraction algorithms have their advantages and disadvantages:

As a frequency domain analysis method, power spectral density can obtain spectrum information intuitively, but it also has an obvious defect, that is, it can't retain phase information, so it can't recover the original signal directly through power spectral density. The recognition capability of the power spectral density can also change with the signal-to-noise ratio (SNR), and is greatly affected by the SNR. As a mature time-frequency analysis method, wavelet transform can recover the original signal, and has the advantages of stable bandwidth and multi-resolution analysis, which is widely used in image signal processing and signal denoising. However, the result of wavelet transform is subject to the selection of wavelet basis function, and there are still some limitations in practical application. As a new method of time-frequency analysis, HHT algorithm can not only accomplish the functions of wavelet transform, but also can obtain the transient parameters without the base function. It is more widely used, but also has the problems such as endpoint effect, mode aliasing effect and negative frequency. The common space pattern is a kind of spatial analysis algorithm, which can effectively maximize the difference between the two kinds of data features, and it is convenient to classify them, but some problems, such as over-fitting and frequency band selection, will also restrict its development.
In practice, power spectral density is often used to observe the distribution and transformation of EEG rhythm. It is a simple and convenient algorithm without recovering the original signal. Wavelet transform is often combined with other algorithms to eliminate interference noise. HHT transform and common space pattern algorithm are widely used and need to be used to denoising and so on. It can be predicted that the future development direction of EEG feature extraction algorithm is to optimize the existing algorithms continuously, and combine two or more algorithms to maximize the advantages of various algorithms. For example, HHT transform was performed after denoising EEG signals with wavelet transform, and CSP treatment was conducted after EEG signals are decomposed into IMF with empirical mode (EMD-CSP algorithm).

With the rapid development of brain-computer interface technology, a lot of ethical issues need to be settled urgently. To find a better algorithm to process EEG signals, to enhance the security of brain-computer interface equipment, and to improve the relevant laws and regulations need the long-term joint efforts of all sectors of society.

References
[1] Ren Yupeng, Wang Guangzhi, Cheng Ming, Gao Xiaorong, Ji Linhong. The Control on the Artificial limb Based on Brain-Computer Interface [J]. Chinese Journal of Rehabilitation Medicine, 2004 (5).
[2] Fan Yingying. Analysis and Reserch on EEG Signals Based on HHT Algorithm [D]. Harbin University of Science and Technology, 2012.
[3] Zhang Xuejun, Wan Dongsheng, Sun Zhixin. EEG signal feature extraction method based on LMD and FE combined with CSP [J]. Chinese Journal of Scientific Instrument, 2020, 41 (8): 227.
[4] RAMOSER H, MULLER-GERJING J, PFURTSCHELLER G. The Optimal Spatial Filtering of EEG in a Single Test During Imaginary Hand Movement [J]. IEEE Rehabilitation Medicine, 2000, 8 (4):441-446.
[5] Zhang Xuejun, Huang Wanlu, Huang Liya, Cheng Xiefeng. EEG Signals Feature Extraction Combined with Empirical Mode Decomposition and Common Spatial Pattern [J]. Computer Engineering and Applications, 2017, 053 (013).
[6] CSDN, 2020. The Characteristics of PSD Power Spectrum Density Extracted From EEG Signals. https://blog.csdn.net/frostime/article/details/106967703?utm_medium=distribute.pc_relevant.none-task-blog-BlogCommendFromBaidu-1.control&depth_1-utm_source=distribute.pc_relevant.none-task-blog-BlogCommendFromBaidu-1.control.
[7] CSDN, 2019. Feature DXtraction Method-Algorithm Principle and Implementation of Common Spatial Pattern (CSP) Algorithm. https://blog.csdn.net/sinat_35162460/article/details/87947521.
[8] J. Mu’ller-Gerking, G. Pfurtscheller, and H. Flyvbjerg. Designing optimal spatial filters for singletrial EEG classification in a movement task. Clinical Neurophysiology, 110(5):787–798, 1999.
[9] S.G. Mason and G.E. Birch. A general framework for brain-computer interface design. Neural Systems and Rehabilitation Engineering, IEEE Transactions on [see also IEEE Trans. on Rehabilitation Engineering], 11(1):70–85, 2003.
[10] D.J. McFarland and J.R. Wolpaw. Sensorimotor rhythm-based brain-computer interface (BCI): feature selection by regression improves performance. Neural Systems and Rehabilitation Engineering, IEEE Transactions on [see also IEEE Trans. on Rehabilitation Engineering], 13(3):372–379, 2005.
[11] B. Blankertz, G. Curiño, and K.R. Mu’ller. Classifying Single Trial EEG: Towards Brain Computer Interfacing. Advances in Neural Information Processing Systems 14: Proceedings of the 2002 [sic]Conference, 2002.
[12] C. Brunner, B. Graimann, J.E. Huggins, S.P. Levine, and G. Pfurtscheller. Phase relationships between different subdural electrode recordings in man. Neuroscience Letters, 375(2):69–74, 2005.

[13] G. Dornhege, B. Blankertz, G. Curio, and K.R. Mu‘ller. Combining features for BCI. Advances in Neural Inf. Proc. Systems (NIPS 02), 15:1115–1122, 2003.

[14] Y. Gao, M.J. Black, E. Bienenstock, S. Shoham, and J.P. Donoghue. Probabilistic Inference of Hand Motion from Neural Activity in Motor Cortex. Advances in Neural Information Processing Systems 14: Proceedings of the 2002 [sic] Conference, 2002.

[15] D.J. McFarland, W.A. Sarnacki, T.M. Vaughan, and J.R. Wolpaw. Brain-computer interface (BCI) operation: signal and noise during early training sessions. Clinical Neurophysiology, 116(1):56–62, 2005.