Multiscale fuzzy entropy based on local mean decomposition and Fisher rule for EEG feature extraction in human motion analysis

Huili He[1,*]

[1]College of Tai Chi Martial Arts, Jiaozuo University, Jiaozuo 454150 China

Abstract

Electroencephalogram (EEG) is a nonlinear, non-stationary, and random weak signal generated by a large number of neurons. It has great research value and practical significance in artificial intelligence, biomedical engineering and other fields. EEG feature extraction is an important step which directly affects the processing results. Currently, the commonly used methods for EEG feature extraction include frequency domain or time domain analysis and time-frequency combination. Due to the nonlinearity of EEG, the above methods have certain limitations. Therefore, this paper proposes a multiscale fuzzy entropy based on local mean decomposition and Fisher rule for EEG feature extraction in human motion analysis. Firstly, the EEG signal is decomposed adaptively into a series of product function (PF) components. Then the effective PF component is selected and the multiscale fuzzy entropy is calculated. Multi-scale fuzzy entropy is used for feature extraction. Fisher rule is used to rank the feature classification ability of fuzzy entropy at different scales, and the multi-scale fuzzy entropy with the highest ranking is selected to form the optimal feature vector to achieve feature dimension reduction. Experimental results show that this proposed method can extract the features of EEG signal effectively, which verifies the validity and feasibility of the new method.

---

[*]Corresponding author. Email: aqiufenga@163.com

1. Introduction

Electroencephalogram (EEG) reflects the functional state of the brain and the electrical activity of brain tissues [1-3]. When performing motor tasks or motor imagination, μ rhythm (8-12Hz) and β rhythm (18-25Hz) responsible for motor perception in the brain will change. An event-related desynchronization (ERD) event or an event-related synchronization (ERS) event will occur. Motor imagination EEG signals are widely used to control brain computer interface (BCI) [4], so feature extraction is the key of brain computer interface technology.

Common algorithms for feature extraction of EEG signal include cospatial mode filtering, autoregressive model and wavelet transform. The common spatial pattern (CSP) algorithm is to filter the EEG signals in spatial domain, so as to extract the EEG characteristics under different motion modes [5,6]. CSP algorithm has achieved good results in binary EEG signals, but it needs to be targeted at a specific frequency band and a large number of electrodes. Adaptive regressive (AR) modeling [7] is used to reflect the time-varying characteristics of EEG signals by using AR models or AR spectrum features. The method is suitable for stationary signal analysis, and EEG signal is a typical non-stationary nonlinear signal. Wavelet Transform (WT) method [8,9] uses variable time-frequency windows to decompose signals step by step, and then selects specific wavelet coefficients as features according to prior

information. However, for EEG signals with complex mechanism, accurate prior information usually cannot be obtained. Local Mean Decomposition (LMD) algorithm [10,11] is an adaptive decomposition of EEG signals to obtain several amplitude-modulated Product Function (PF) components, which can reflect the time-frequency changes of signals [12]. Commonly used entropy analysis includes approximate entropy (AE), sample entropy (SE), fuzzy entropy (FE) and multi-scale fuzzy entropy (MFE). The fuzzy entropy [13,14] feature has clear physical significance and can measure the probability of generating new patterns. It replaces the binarization function in sample entropy with continuous exponential function as similarity measure and overcomes the mutation problem in entropy calculation. In view of this, Torrents-Barrena et al. [15] used fuzzy entropy to diagnose Alzheimer's disease and achieved good accuracy. In reference [16], normalized brainwave power gain and fuzzy entropy were used as features to construct a prediction model for drunk driving accidents, and the experimental results showed that the estimated value of the model was consistent with the real value. Therefore, fuzzy can be used to analyze non-stationary and non-stationary EEG signals. According to different motion picture mode nonlinear non-stationary EEG feature extraction problem, this research combines LMD and multi-scale fuzzy entropy (MSE) for imagine EEG feature extraction, the original EEG signals are decomposed, the calculation of the effective product function (PF) component of multi-scale entropy is as a feature vector, then it used support vector machine (SVM) for classification and recognition.

2. Proposed EEG feature extraction method

The local mean decomposition (LMD) method was proposed in 2015 [17]. It was first applied in feature extraction of EEG signals, and then it was widely used in other fields, such as mechanical fault diagnosis. In signal feature extraction, the time-domain statistics of PF component of LMD are often extracted as features, which are easily affected by the noise in motor EEG signal and can not extract effective feature vectors.

2.1. Local mean decomposition(LMD)

Local mean decomposition is an adaptive time-frequency analysis method for nonstationary nonlinear signals. The essence of LMD is to adaptively decompose EEG signals into multiple product function (PF) components and a residual quantity R. Each PF component is the product of a pure FM signal and an envelope signal. The instantaneous frequency of the PF component can be calculated from the FM signal, and the instantaneous amplitude of the PF component can be obtained from the envelope signal. For the EEG signal $x(t)$, its specific decomposition process is as follows:

Find all local maximum points and minimum points $n_i (i = 1,2,\cdots)$ of EEG signal $x(t)$. Find the mean value $m_i = (n_i + n_{i+1})/2$ of two adjacent extremum points n_i and n_{i+1}. The local mean function $m_{11}(t)$ is obtained by connecting adjacent mean points m_i with straight lines and smoothing them with moving average method.

(2) The envelope estimation value $a_i = (|n_i + n_{i+1}|)/2$ is obtained by using the local mean point n_i. The adjacent envelope estimation points a_i are connected by straight lines, and then the envelope estimation curve $a_{11}(t)$ is obtained by smoothing with the moving average method.

(3) The local mean function $m_{11}(t)$ is separated from the original signal $x(t)$, and $h_{11}(t) = x(t) - m_{11}(t)$ is obtained.

(4) Divide the separated $h_{11}(t)$ by $a_{11}(t)$ for demodulation and get $s_{11}(t) = h_{11}(t)/a_{11}(t)$. Determine whether $s_{11}(t)$ is a pure frequency modulation function, that is, whether the envelop function $a_{12}(t)$ of $s_{11}(t)$ is equal to 1. If $a_{12}(t) = 1$, $s_{11}(t)$ is a pure frequency modulation function. If $a_{12}(t) \neq 1$, continue to repeat the above steps k times and get $s_{1k}(t)$, until the envelope function meets $a_{1(k+1)}(t) = 1$, namely, $h_{1n}(t) = s_{1(k-1)}(t) - m_{1k}(t)$. The iteration termination condition is $\lim_{n \to \infty} a_{1k}(t) = 1$.

(5) The envelope of the first component $a_1(t) = a_{11}(t)a_{12}(t)\cdots a_{1k}(t) = \prod_{d=1}^{k} a_d(t)$ is obtained by multiplying the obtained envelope estimation function $a_{1d}(t)$.

(6) The first PF component of the original signal is the envelope signal $a_1(t)$ and pure FM signal $s_{1k}(t)$, that is, $PF_1 = a_1(t)s_{1k}(t)$.

(7) $PF_1(t)$ is separated from the original signal, and the new signal $u_1(t)$ is used as the source signal to repeat the above steps for q times until $u_q(t)$ is a monotone function and the iteration equation is $u_q(t) = u_{q-1}(t) - PF_q(t)$. Finally, the original signal $x(t)$ is expressed as

$$x(t) = \sum_{j=1}^{q} PF(t) + u_q(t) \quad . \quad u_q(t) \text{ is the residual}$$

component.

After the decomposition of EEG signal by LMD, the characteristic information of the original signal will be distributed to different time characteristic scales, thus amplifying the hidden characteristic information of the

original signal. Then multi-scale fuzzy entropy can be used to quantitatively evaluate the rationality of PF components more easily.

## 2.2. Multiscale fuzzy entropy (MFE)

The multi-scale fuzzy entropy algorithm is composed of improved coarse-granulating process and fuzzy entropy algorithm [18-20].

Moving mean filtering coarsening

The traditional coarse-granulating process of time series coarse-granulates the original series with length N at various scales [21]. The coarse-granulating sequence is:

$$y_i(k) = \frac{1}{k} \sum_{i=(j-1)k+1}^{jk} x_i \qquad (1)$$

Where $k = 1,2,3$ is the scale factor, $1 \le j \le N/k$. According to the formula (1), the sequence length after traditional coarse-granulating becomes 1/k of the original sequence length. When k value is large, the data length cannot meet the requirements of fuzzy entropy calculation, and the coarse-granulating of different scale factors leads to inconsistent data shrinkage. The simplified average processing of equation (1) is also easy to cause information loss. On the other hand, when the sequence length is not a multiple of k, some information will be ignored, which will affect the accuracy and stability of the multi-scale entropy algorithm. In view of the above shortcomings, this paper adopts the moving mean filtering coarse-granulating algorithm, taking scale 3 as an example, and its calculation process is shown in figure 1.
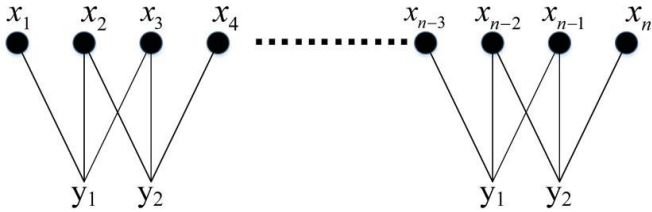
Figure 1. Coarse-granulating process of moving average

The sequence length after coarse-granulating by the moving mean filter is N-k+1, and the time series after coarse-granulating is expressed as:

$$y_j(k) = \frac{1}{k} \sum_{i=j}^{j+k-1} x_i, (1 \le j \le N-k+1) \qquad (2)$$

The moving average method reduces the dependence on the length of original time series, avoids data loss and improves the accuracy of feature extraction.

Fuzzy entropy

The similarity measure formula of fuzzy entropy uses exponential function as fuzzy function and the continuity of exponential function makes fuzzy entropy smooth [22,23]. In addition, fuzzy entropy introduces the concept of fuzzy set to measure the similarity of two vectors. Fuzzy entropy is defined as follows:

1) Suppose there is a time series $\{u(i):1 \le i \le N\}$ with length N. Regenerate a set of m-dimensional vectors in sequential order.

$$X_i^m = \{u(i), u(i+1), \cdots, u(i+m-1)\} - u_0(i) \quad (3)$$

Where $u_0(i)$ is the mean value of m values.

2) Define the maximum Euclidean distance d_{ij}^m of X_i^m and X_i^m.

$$d_{ij}^m = d[X_i^m, X_i^m] = \max_{k \in (0, m-1)} \{|(u(i+k) - u_0(i)) - u(j+k) - u_0(j))|\} \qquad (4)$$

Where $i, j = 1, 2, \cdots, N = m, i \ne j$.

3) Define the similarity D_{ij}^m of vector X_i^m and X_i^m by fuzzy function $\mu(d_{ij}^m, n, r)$, i.e.,

$$D_{ij}^m = \mu(d_{ij}^m, n, r) = e^{-(d_{ij}^m)^n} \qquad (5)$$

Where, n and r are the boundary gradient and width of the fuzzy function respectively.

4) Defining function

$$\varphi^m(n, r) = \frac{1}{N-m} \cdot \sum_{i=1}^{N-m} (\frac{1}{N-m-1} \sum_{j=1, j \ne i}^{N-m} D_{ij}^m) \qquad (6)$$

5) Similarly, repeat steps (2)~(4) for dimension (m+1) and get φ^{m+1}.

6) Defining fuzzy entropy

$$Fuzzy\ En(m, n, r) = \lim_{N \to \infty} [\ln \varphi^m(n, r) - \ln \varphi^{m+1}(n, r)] \quad (7)$$

When the amount of data is limited, the above equation can be approximately expressed as:

$$Fuzzy\ En(m, n, r, N) = \ln \varphi^m(n, r) - \ln \varphi^{m+1}(n, r) \qquad (8)$$

Obviously, it can be seen from equation (7) and coarse-granulating process that the preset parameters required for multi-scale fuzzy entropy calculation include embedding dimension m, similarity tolerance r and scale factor k. According to abundant experiments, m=2, r=0.2SD(SD is the standard deviation of the original sequence) and k=20 are selected in this paper.

2.3. Feature selection based on Fisher score

After feature extraction by multi-scale fuzzy entropy algorithm, multiple feature vectors are obtained. However, not all feature vectors are closely related to task classification. Features also contain too much redundant information, even noise information, which affects the accuracy of classification. In order to reduce the feature dimension and improve the classification accuracy, Fisher

score is used to screen the features. The basic idea of Fisher score [24,25] is to calculate the inter-class variance and intra-class variance ratio of features according to Fisher criterion. Fisher scores for the two types of samples are defined as follows:

There exists a sample set $\{m_1, m_2, \cdots, m_i, \cdots, m_k\}$, k is the sample number. So let the left hand data set be L, and the number is l, and the right hand data set be R, and the number is r. The Fisher score corresponding to the n-th feature is defined as:

$$F_n = S_{b,n} / S_{w,n} \tag{9}$$

In here,

$$S_{b,n} = (\mu_{1,n} - \mu_n)^2 + (\mu_{2,n} - \mu_n)^2 \tag{10}$$

$\mu_{1,n}$, $\mu_{2,n}$, and μ_n represent the mean values of imagining left hand, imagining right hand, all samples in the n-th dimension, respectively. $S_{b,n}$ represents the inter-class variance of the n-th feature on the training sample set, and describes the distance between samples of different classes.

$$S_{w,n} = \frac{1}{l} \sum_{x \in L} (m_n - \mu_{1,n})^2 + \frac{1}{r} \sum_{x \in R} (m_n - \mu_{2,n})^2 \tag{11}$$

m_n is the value of sample m in the n-th dimension. $S_{w,n}$ represents the in-class variance of the n-th feature on the training sample set, and describes the distance between similar samples. A larger Fisher score, namely, the intra-class distance is small and the inter-class distance is large, which indicates that this feature has a relatively large degree of differentiation between categories and a good classification ability. As shown in reference [26], seven features with the highest score were selected to distinguish Alzheimer's disease from normal people. In reference [27], the optimal features were selected by adding features in sequence and combining them with support vector machine according to the score, and the optimal feature subset was found in all feature vectors. Although this method could obtain the global optimal feature subset, Fisher score indicated that the feature of feature classification ability was not fully utilized. Therefore, this paper selects the five features with the highest score of each channel and combines them with support vector machine to select the optimal feature vector.

2.4. Combining LMD and MFE

Different motor imagery patterns trigger electrical activity in different areas of the cerebral cortex, and distributed electrodes in the EEG acquisition system record changes in the electrical signals in the brain. When imagining unilateral limb movement, the amplitude of spectrum oscillation of rhythm μ and rhythm β in contralateral motor sensory regions of the brain is reduced or blocked as event-related synchronous events. The electrical activity events of μ rhythm and β rhythm increase the amplitude of the

spectrum de-correlated and de-synchronized events. The μ rhythm is concentrated in 8-12Hz. β The rhythm is concentrated in 18-25 Hz. Generally, EEG signals with 8~30Hz are used for feature extraction of these two rhythms.

Aiming at the nonlinear and non-stationary characteristics of motor imagery EEG signals, a feature extraction method based on LMD and MFE is proposed in this paper. In the decomposition process of LMD, the signals are decomposed step by step to retain the essential features of the original signal. The characteristic information of the original signal is displayed at different resolutions. It is easier to extract the characteristic information through multiple PF components. MSE method is an effective method to describe information features, which can quantitatively describe EEG signals by calculating multi-scale entropy of PF components. The combination of the two methods can effectively analyze and extract the features of motor imagination EEG signals, which is conducive to the recognition of motor imagination categories.

The feature extraction process of motor imagination EEG signals is shown in figure 2. The specific realization steps of feature extraction method of motor imagery EEG based on LMD and MFE are as follows.
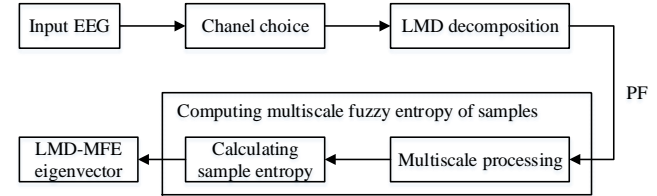

Figure 2. EEG feature extraction process

(1) Input EEG signals and perform LMD decomposition on each sample to obtain a series of PF components.
(2) The EEG signals are decomposed successively, and the feature information is mainly distributed in the first few PF components, so the first few PF components are selected to extract the feature information.
(3) Calculate the multi-scale entropy values of PF components of each training sample and form them into feature vectors.
(4) The feature vector is used as the input of SVM classifier to train the classifier.
(5) The feature vectors of the samples to be tested are input into the trained SVM classifier, and the classification of motion imagination is determined by the output of SVM classifier.

2.5. Feature recognition by support vector machine

Support vector machine (SVM) is an excellent classification algorithm, which is mainly based on the structural risk minimization theory of statistical learning theory and VC (Vapnik-Chervonenkis) dimension [28-30]. The support

vector machine transforms the input sample into a higher dimensional space through a mapping, and then finds the optimal classification surface in this higher dimensional space, thus separating the samples.
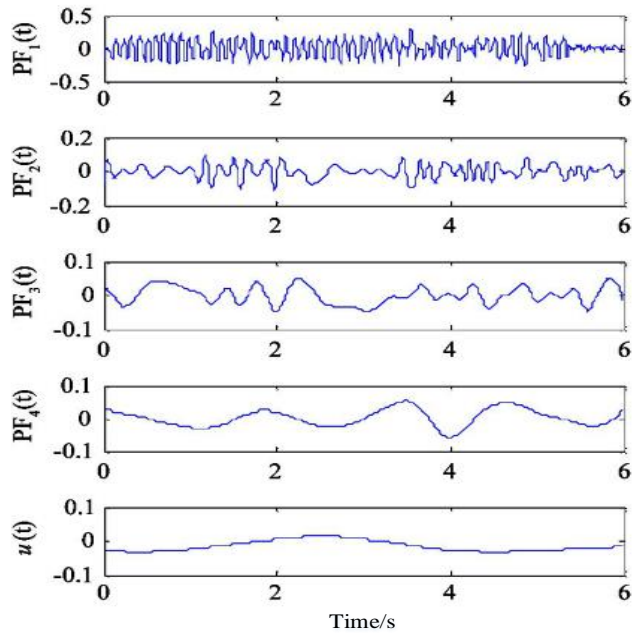
If the training set sample is a d-dimensional vector, it has n samples. The category of each sample is expressed as $(x_i, y_i), y_i \in \{-1,1\}, i = 1,2,\cdots, n$. In d-dimensional space, the linear discriminant function is $g(x) = w \cdot x + b$, then the hyperplane equation is $w \cdot x + b = 0$. The discriminant function is normalized to make all samples of the two classes satisfy $y_i(wx_i + b) \geq 1$ in the case of linear indivisibility, add the relaxation term $\eta_i \geq 0$ into the equation at the same time, then the objective function is:

$$\min \frac{1}{2} \| w \|^2 + C(\sum_{i=1}^{n} \eta_i) \qquad (12)$$

Where w is the hyperplane normal vector and the constant C is the penalty factor. For test set data, its discriminant function is:

$$f(x) = \text{sgn}(\sum_{i=1}^{n_s} a_i^* y_i K(x_i, x) + b) \qquad (13)$$

In the formula, $sign(\cdot)$ is the sign function, and the vector x_i corresponding to a_i^* is the support vector. n_s is the number of support vectors, $K(x_i, x)$ is the kernel, and b is the intercept.

3. Experiments analysis

This paper uses dataset III [31] to verify the proposed algorithm. The experimental data records the EEG signals of a 25-year-old normal female in imaginary right-handed movement. The experiment is conducted on the same day, including seven groups with 40 times in each group. The EEG data of channel C3, C4 and Cz are recorded with a sampling frequency 128Hz and band-pass filtering frequency 0.5-30Hz. Each experiment lasts for 9s, 0~2s is the resting state. The experiment begins with a sound reminder from 2s, and a "+" is displayed on the screen at the same time, indicating the end of 1s. From the third second, the arrow on the screen prompts the experimenter to imagine movement, and the end is 9s.

The data of C3 and C4 channels are selected for feature extraction. The time of motion imagination is 3-9s, so the data between 3-9s of each channel are extracted, and each channel of each group has a total of 128×6=768 data points, and then the data of each channel is decomposed by LMD adaptive method. Figure 3 is the time-frequency diagram of LMD decomposition of C4 channel when imagining right hand motion.
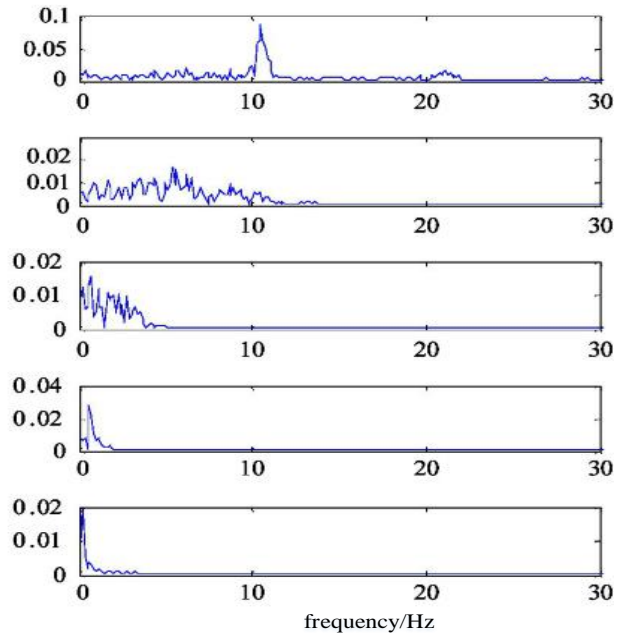
Figure 3. Time-frequency diagram of each PF component of EEG signal (C4 channel) decomposed by LMD

As can be seen from the left column of figure 3, EEG signals are adaptively decomposed into four PF components and one residual component u(t). The spectrum of PF and residual components is shown in the right column of figure 3. It is obvious from the figure that there are large fluctuations around 10Hz and 20Hz, indicating the event dependent synchronization/ event dependent dessynchronization (ERS/ERD) phenomenon [32,33]. The experimental analysis shows that μ rhythm and β rhythm information including the left-right motion imagination are mainly distributed in the first three PF components. Therefore, the first three PF components are selected as the signal of feature extraction.

The time of motion imagination is 6s. Considering the processing time and classification accuracy, the time period used to extract the optimal feature is selected. The sliding time window is 2s and the sliding step is 1s. The results

show that the classification effect is better in 4-6s and 5-7s, so 4-7s is selected as the feature extraction data segment.

The similarity tolerance r=0.35SD (SD is the standard deviation of the original signal) is selected according to several experiments, and the entropy values of multiple EEG signals at different scales are calculated, so as to select the parameters of multi-scale entropy. As shown in figure 4, the entropy of PF1 and PF2 increase first and then decrease with the increase of scale factor τ, and the entropy of PF3 increase with the increase of scale factor τ. When the experimental scale factor τ =9, the entropy values of each component are discriminated and the classification accuracy is high. Therefore, τ =9 is selected. Table 1 shows the average multi-scale entropy values of C3 and C4 channels of different categories when r= 0.35SD, τ =9 is used to imagine the right hand motion.
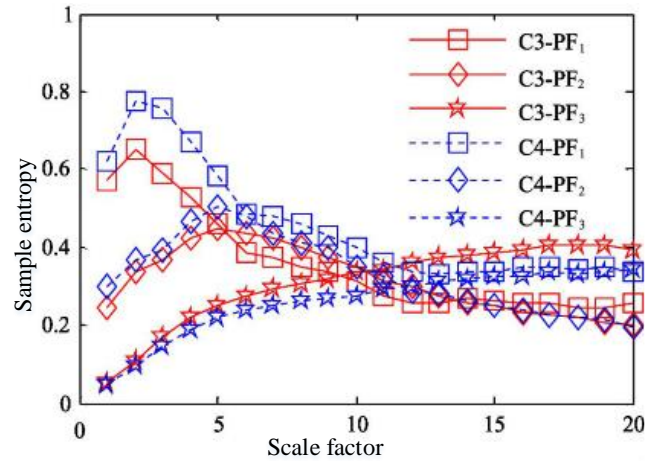
(b) The change of multiscale fuzzy entropy of the imagine right hand movement
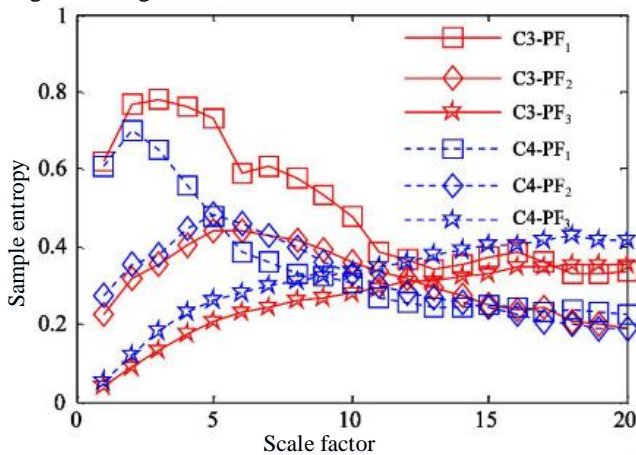

(a) The change of multiscale fuzzy entropy of the imagine left hand movement

Table 1. The mean multiscale fuzzy entropy of Imagine right hand motion

Scale factor	C3			C4		
	PF1	PF2	PF3	PF1	PF2	PF3
3	0.5996	0.3771	0.1781	0.7644	0.4044	0.1555
6	0.3954	0.4453	0.2858	0.4960	0.4887	0.2498
9	0.3467	0.3821	0.3286	0.4429	0.4078	0.2816
12	0.2657	0.3085	0.3707	0.3447	0.2997	0.3126
15	0.2711	0.2602	0.3999	0.3486	0.2594	0.3379
18	0.2547	0.2276	0.4165	0.3518	0.2293	0.3443

The first PF component of C3 and C4 channels is selected to calculate the multi-scale fuzzy entropy of the above conditions, as shown in figure 5. As can be seen from figure 5, although the multi-scale fuzzy entropy of EEG signals imagining left-right movement overlaps partially, it still has a degree of differentiation. Therefore, multi-scale fuzzy entropy can reflect the characteristics of EEG signals in imagining left-right movement. The multi-scale fuzzy entropy of the first three order PF components contributes the most to the classification. The combination of the first three order multi-scale fuzzy entropy as feature vector is expected to improve the classification accuracy.
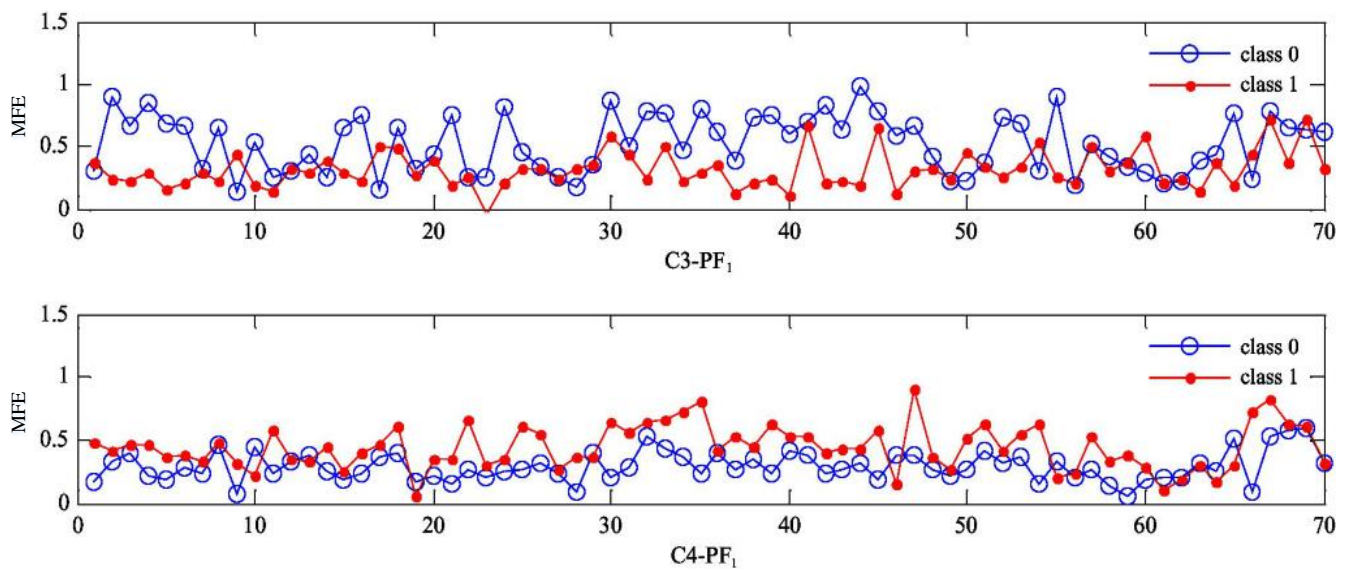
Figure 5. Multiscale fuzzy entropy distribution

Next, support vector machine classifier is used to classify the extracted feature vectors. Support vector machine (SVM) is a machine learning method based on statistical learning theory. It has many advantages in solving small sample classification, nonlinear problems and high-dimensional pattern recognition. In order to accelerate the convergence of the training network, the training set and test set are normalized first, in order to summarize the statistical distribution of uniform data. Finally, it inputs the feature vectors of the test set for classification and outputs the motion imagination category.

The experimental dataset consists of 280 trials, which are randomly divided into two groups with 140 trials each. The 6-dimensional feature vectors extracted by multi-scale entropy are input into SVM classifier for training classification, and the optimal classification recognition rate is 85.21%. As can be seen from table 2, the classification accuracy of the feature extraction algorithm in this paper is 0.92%-4.13% higher than that in references SDA [34], EPCA [35] and NVDNN [36]. In this paper, 6-dimensional feature vectors are used. The reduction of feature dimension can make the classifier model simpler and reduce the classification time. Compared with the existing literatures, the number of features is significantly reduced, and 100 tests are carried out. The average time is 73ms. Thus, the new algorithm in this paper reduces the number of features and improves the classification accuracy.

Table 2. Recognition rate under different feature extraction methods

Feature	Dimension	Accuracy/%
SDA	14	81.08
EPCA	8	82.86
NVDNN	8	84.29
Proposed	6	85.21

4. Conclusion

A feature extraction method based on LMD and multi-scale fuzzy entropy is proposed in this paper. LMD is used to decompose the EEG signals of left and right motor imagery adaptively, and the decomposed PF components are extracted from the feature vectors of multi-scale entropy, which are input into SVM for classification and recognition, thus realizing the classification of motor imagery EEG signals. Experimental results show that LMD and multi-scale fuzzy entropy are better than traditional feature extraction methods in recognition of EEG signals. The results show that the proposed method can effectively extract features from motor imagery EEG signals

References

[1] Ahmad F K. A Comparative Analysis of Time-frequency Feature Extraction Techniques for Large Scale Electroencephalogram Data[J]. International Journal of Advanced Trends in Computer Science and Engineering, 2021, 10(1):14-24.

[2] Li H, Yin S, Liu J, et al. Novel gaussian approximate filter method for stochastic non-linear system[J]. International Journal of Innovative Computing, Information and Control. 13(1): 201-218, 2017.

[3] Shoulin Yin, Ye Zhang, Shahid Karim. Large Scale Remote Sensing Image Segmentation Based on Fuzzy Region Competition and Gaussian Mixture Model[J]. IEEE Access. volume 6, pp: 26069 - 26080, 2018.

[4] Bashar S K, Bhuiyan M. Classification of motor imagery movements using multivariate empirical mode decomposition and short time Fourier transform based hybrid method[J]. Engineering Science & Technology An International Journal, 2016:S2215098616302592.

[5] F. A. Alturki, M. Aljalal, A. M. Abdurraqeeb, K. Alsharabi and A. A. Al-Shamma'a, "Common Spatial Pattern Technique With EEG Signals for Diagnosis of Autism and Epilepsy Disorders," in IEEE Access, vol. 9, pp. 24334-24349, 2021, doi: 10.1109/ACCESS.2021.3056619.

[6] A. Jiang, Q. Wang, J. Shang and X. Liu, "Sparse Common Spatial Pattern for EEG Channel Reduction in Brain-Computer Interfaces," 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP), 2018, pp. 1-4, doi: 10.1109/ICDSP.2018.8631618.

[7] U. Rashid, I. K. Niazi, M. Jochumsen, L. R. Krol, N. Signal and D. Taylor, "Automated Labeling of Movement-Related Cortical Potentials Using Segmented Regression," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 27, no. 6, pp. 1282-1291, June 2019, doi: 10.1109/TNSRE.2019.2913880.

[8] Taran S, Bajaj V. Motor imagery tasks-based EEG signals classification using tunable-Q wavelet transform[J]. Neural Computing and Applications, 2019, 31(11):6925-6932.

[9] Yin Shoulin, Liu Jie, Teng Lin. A new krill herd algorithm based on SVM method for road feature extraction[J]. Journal of Information Hiding and Multimedia Signal Processing, v 9, n 4, p 997-1005, July 2018.

[10] Deng L, Zhao R. Fault feature extraction of a rotor system based on local mean decomposition and Teager energy kurtosis[J]. Journal of Mechanical Science & Technology, 2014, 28(4):1161-1169.

[11] Zhang T, Chen W. LMD based features for the automatic seizure detection of EEG signals using SVM[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2017:1100-1108.

[12] David J R, Jatkar D P, Sen A. Product Representation of Dyon Partition Function in CHL Models[J]. Journal of High Energy Physics, 2006, 2006(6).

[13] Li M, Wang R, Xu D. An Improved Composite Multiscale Fuzzy Entropy for Feature Extraction of MI-EEG[J]. Entropy, 2020, 22(12):1356.

[14] Huang W, Wu M, Chen L, et al. Multi-objective drilling trajectory optimization using decomposition method with minimum fuzzy entropy-based comprehensive evaluation[J]. Applied Soft Computing, 2021, 107(12):107392.

[15] Torrents-Barrena J, Puig D, Rajeesh J, et al. A Novel Fuzzy-Entropy Threshold based on a Complex Wavelet Denoising Technique to Diagnose Alzheimer Disease[J]. Healthcare Technology Letters, 2016.

[16] N. Fuad, W. Mansor and K. Y. Lee, "Wavelet packet analysis of EEG signals from dyslexic children with writing disability," 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013, pp. 2890-2893, doi: 10.1109/EMBC.2013.6610144.

[17] Li Y, Xu M, Haiyang Z, et al. A new rotating machinery fault diagnosis method based on improved local mean decomposition[J]. Digital Signal Processing, 2015, 46(C):201-214.

[18] Li Y, Xu M, Wang R, et al. A fault diagnosis scheme for rolling bearing based on local mean decomposition and improved multiscale fuzzy entropy[J]. Journal of Sound and Vibration, 2016, 360:277-299.

[19] Azami H, Escudero J. Refined Composite Multivariate Generalized Multiscale Fuzzy Entropy: A Tool for Complexity Analysis of Multichannel Signals[J]. Physica A: Statistical Mechanics and its Applications, 2016, 465.

[20] Z. Cao and C. Lin, "Inherent Fuzzy Entropy for the Improvement of EEG Complexity Evaluation," in IEEE Transactions on Fuzzy Systems, vol. 26, no. 2, pp. 1032-1035, April 2018, doi: 10.1109/TFUZZ.2017.2666789.

[21] S. Dick, A. Schencker, M. Last, H. Bunke and A. Kandel, "Re-granulating a fuzzy rulebase," 10th IEEE International Conference on Fuzzy Systems. 2001, pp. 372-375 vol.1, doi: 10.1109/FUZZ.2001.1007326.

[22] Teng Lin, Hang Li and Shoulin Yin. Modified Pyramid Dual Tree Direction Filter-based Image De-noising via Curvature Scale and Non-local mean multi-Grade remnant multi-Grade Remnant Filter [J]. International Journal of Communication Systems. Vol. 31, no. 16, November 10,pp. e.3486.1-e.3486.12, 2018.

[23] Shoulin Yin, Ye Zhang and Shahid Karim. Region search based on hybrid convolutional neural network in optical remote sensing images[J]. International Journal of Distributed Sensor Networks, Vol. 15, No. 5, 2019. DOI: 10.1177/1550147719852036

[24] Lin Teng, Hang Li, Shoulin Yin, Yang Sun. Improved krill group-based region growing algorithm for image segmentation[J]. International Journal of Image and Data Fusion. 10(4), pp. 327-341, 2019. doi: 10.1080/19479832.2019.1604574

[25] Lin Teng, Hang Li, Shoulin Yin, Shahid Karim &Yang Sun. An active contour model based on hybrid energy and fisher criterion for image segmentation[J]. International Journal of Image and Data Fusion. Vol.11, No. 1, pp. 97-112. 2020.

[26] Khazaee A, Ebrahimzadeh A, Babajani-Feremi A. Automatic classification of Alzheimer's disease with resting-state fMRI and graph theory. IEEE, 2015.

[27] Zhang RL, Zhang Z, Peng XJ, et al. Feature selection algorithm based on Fisher value and support vector machine[J]. Computer Engineering and Design, 2014, 35 (12): 4145-4148.

[28] Jing Yu, Hang Li, Shoulin Yin. Dynamic Gesture Recognition Based on Deep Learning in Human-to-Computer Interfaces [J]. Journal of Applied Science and Engineering, vol. 23, no. 1, pp.31-38, 2020.

[29] Shoulin Yin, Hang Li, Lin Teng, et al. An optimised multi-scale fusion method for airport detection in large-scale optical remote sensing images [J]. International Journal of Image and Data Fusion, vol. 11, no. 2, pp. 201-214, 2020. DOI: 10.1080/19479832.2020.1727573

[30] Yin, S., Li, H. GSAPSO-MQC:medical image encryption based on genetic simulated annealing particle swarm optimization and modified quantum chaos system. Evolutionary Intelligence (2020). doi: 10.1007/s12065-020-00440-6

[31] B. Blankertz et al., "The BCI competition 2003: progress and perspectives in detection and discrimination of EEG single trials," in IEEE Transactions on Biomedical

Engineering, vol. 51, no. 6, pp. 1044-1051, June 2004, doi: 10.1109/TBME.2004.826692.

[32] Shoulin Yin, Hang Li, Desheng Liu and Shahid Karim. Active Contour Modal Based on Density-oriented BIRCH Clustering Method for Medical Image Segmentation [J]. Multimedia Tools and Applications. Vol. 79, pp. 31049-31068, 2020.

[33] S. Yin and H. Li. Hot Region Selection Based on Selective Search and Modified Fuzzy C-Means in Remote Sensing Images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 5862-5871, 2020, doi: 10.1109/JSTARS.2020.3025582.

[34] Yu Z, Li L, Zhang W, et al. An Adaptive EEG Feature Extraction Method Based on Stacked Denoising Autoencoder for Mental Fatigue Connectivity[J]. Neural Plasticity, 2021, 2021(5):1-12.

[35] L. Cheng, D. Li, G. Yu, Z. Zhang, X. Li and S. Yu, "A Motor Imagery EEG Feature Extraction Method Based on Energy Principal Component Analysis and Deep Belief Networks," in IEEE Access, vol. 8, pp. 21453-21472, 2020, doi: 10.1109/ACCESS.2020.2969054.

[36] Mouleeshuwarapprabu R, Kasthuri N. Nonlinear vector decomposed neural network based EEG signal feature extraction and detection of seizure[J]. Microprocessors and Microsystems, 2020, 76(4):103075.