ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Based on multi-dimensional signal processing technology for auditory synesthesia music emotions

Xiaohong Cui[1,*], *Xiaoqing Li*[1] *and Dan Mao*[1]

[1] School of Music and Dance, Chongqing Preschool Education College, Chongqing, 404047, China

SUMMARY: *The use of multi-dimensional information technology to process auditory synesthesia in musical emotions has gradually become a hot topic of interest among many scholars. This paper proposes a novel method for decoding electroencephalography (EEG) signals by combining spectral summation analysis (SSA) with entropy measures. EEG time series are decomposed and reconstructed using SSA to obtain SSA components of various orders. Four entropy measure feature data—approximate entropy, sample entropy, fuzzy entropy, and multiscale entropy—are extracted from the SSA components. These entropy measure feature data are then used to construct a feature vector describing the relevant EEG data decoding recognition, and combined with a pattern classifier to achieve EEG decoding. Additionally, a TCM-CSP emotional EEG analysis method based on cognitive topological constraints is proposed. Considering the temporal changes in emotions, an emotional brain region analysis method is proposed, and combined with the spatial characteristics of EEG signals, a CSP computation method that preserves topological constraints is proposed. Finally, experiments show that in emotional tendency analysis, the distortion group and the reverb group exhibit negative correlations. The distortion group tends toward restlessness and anxiety, while the reverb group tends toward calmness and relaxation, indicating that the tonal quality of sound after different effects processing significantly influences emotional perception.*

KEYWORDS: *multidimensional information technology; EEG topological constraints; EEG signals; musical emotion*

1 Introduction

Music is one of the most important forms of human emotional expression, possessing the ability to convey and express emotions, thereby evoking rich emotional experiences in listeners [1, 2]. Music emotion analysis refers to the process of extracting and analyzing emotional information embedded in music through scientific methods, with the aim of identifying and understanding the emotional content of music [3, 4]. Music emotion analysis is a complex and challenging task, as emotions in music are intertwined with various elements such as sound, rhythm, and melody [5-7]. Traditional rule-based music emotion analysis methods often rely on expert knowledge and manually defined rules, making them difficult to generalize across different music genres and cultural contexts [8-10]. Music emotion analysis methods based on multi-dimensional signal processing technology can better capture and analyze the emotions in music by learning the advanced features and patterns of audio signals [11-13].

Multidimensional signal processing technology refers to the process of processing analog

*cuixiaohong888888@163.com
https://doi.org/10.65102/is2026008 Published: 30 April, 2026

signals using digital computing technology. The primary task of digital signal processing is to digitize and algorithmically process complex analog signals to better meet the needs of modern communication, computer, and control fields [14-17]. Signal processing technology has widespread applications, including audio processing, image processing, acoustic signal processing, video processing, measurement, and control fields [18, 19]. In the field of audio processing, multidimensional signal processing technology can process sound through techniques such as filtering, equalization, compression, and distortion, analyze auditory and synesthetic perceptions to achieve emotional analysis of music, and offers advantages over traditional analog signal processing in terms of efficient, precise, and reliable signal processing [20-23].

The study first briefly introduces the main technical methods for decoding psychological states based on EEG, analyzes the factors influencing the performance of existing EEG decoding algorithms based on entropy measures, and discusses a method for extracting EEG entropy features by integrating time series decomposition and signal reconstruction methods with entropy measures. It proposes an EEG signal decoding algorithm combining SSA and entropy measures. Next, a CNN-based CSP feature analysis method is proposed, utilizing a CNN network for feature matrix learning. By defining an optimization criterion for the fully connected layer weight matrix, a dimension-reduced, efficient feature set is obtained. Furthermore, emotional space cognitive patterns are incorporated into the CSP computation method, proposing a CSP computation method based on topological constraint preservation. Based on the distribution differences of emotional cognitive brain regions in the human brain, an emotional topological constraint matrix quantification index (TCM) is established and applied to the extraction of emotion-related EEG features. Subsequently, an auditory synesthesia music emotion experiment is conducted. EEG data and self-reported emotion regulation data are collected through EEG experiments and psychological experimental methods. After data collection and experimental analysis, the phase synchronization index was analyzed to explore the synchrony between auditory synesthesia and music emotion, and the cognitive mechanisms of the brain for different emotions were investigated.

2 Brain signal decoding method based on entropy measure

Entropy is a key concept in the field of nonlinear science, serving as a measure of uncertainty or complexity in dynamical systems. In information theory, entropy can be used to quantify information, thereby serving as a quantitative indicator of a system's information content. Research on EEG signal decoding based on entropy measures primarily involves two processes: EEG entropy feature extraction and machine learning of EEG patterns. The EEG entropy feature extraction process primarily involves using various entropy measurement feature methods to extract EEG entropy feature data from EEG data corresponding to the decoding task. The machine learning process for EEG patterns primarily involves using machine learning algorithms to build EEG pattern classification and recognition models, and using EEG entropy measurement features as input feature vectors to achieve recognition and prediction of EEG decoding tasks.

This paper proposes a brainwave signal decoding method based on the combination of singular spectrum analysis and entropy measures. Singular spectrum analysis is a powerful method for nonlinear time series analysis, whose basic principle is based on principal component analysis of time series [24]. It starts from the dynamical analysis of time series, constructs the trajectory matrix of the time series using phase space reconstruction, and through singular value decomposition and diagonalization averaging of the trajectory matrix, can effectively decompose the time series into a series of important component structures. EEG

signals typically exhibit characteristics such as non-stationarity, nonlinearity, and randomness. The SSA method has a significant advantage in processing this type of signal.

3 Brain signal decoding method based on a combination of SSA and entropy measure

3.1 System block diagram of the proposed brain signal decoding algorithm

Electroencephalographic (EEG) signals are typically non-stationary, nonlinear random signals that generally contain multiple component structures and noise interference. The Singular Spectrum Analysis (SSA) method is a powerful technique for analyzing nonlinear time series data, capable of reconstructing a series of component structures within the time series and effectively extracting time series information from noisy signals [25]. The EEG signal decoding algorithm proposed in this chapter integrates the signal processing method of singular spectrum analysis into the entropy feature extraction process of EEG signals, designing a novel EEG decoding algorithm that combines entropy measures with singular spectrum analysis. The framework of the EEG signal decoding algorithm based on the combination of SSA and entropy measures is shown in Figure 1.

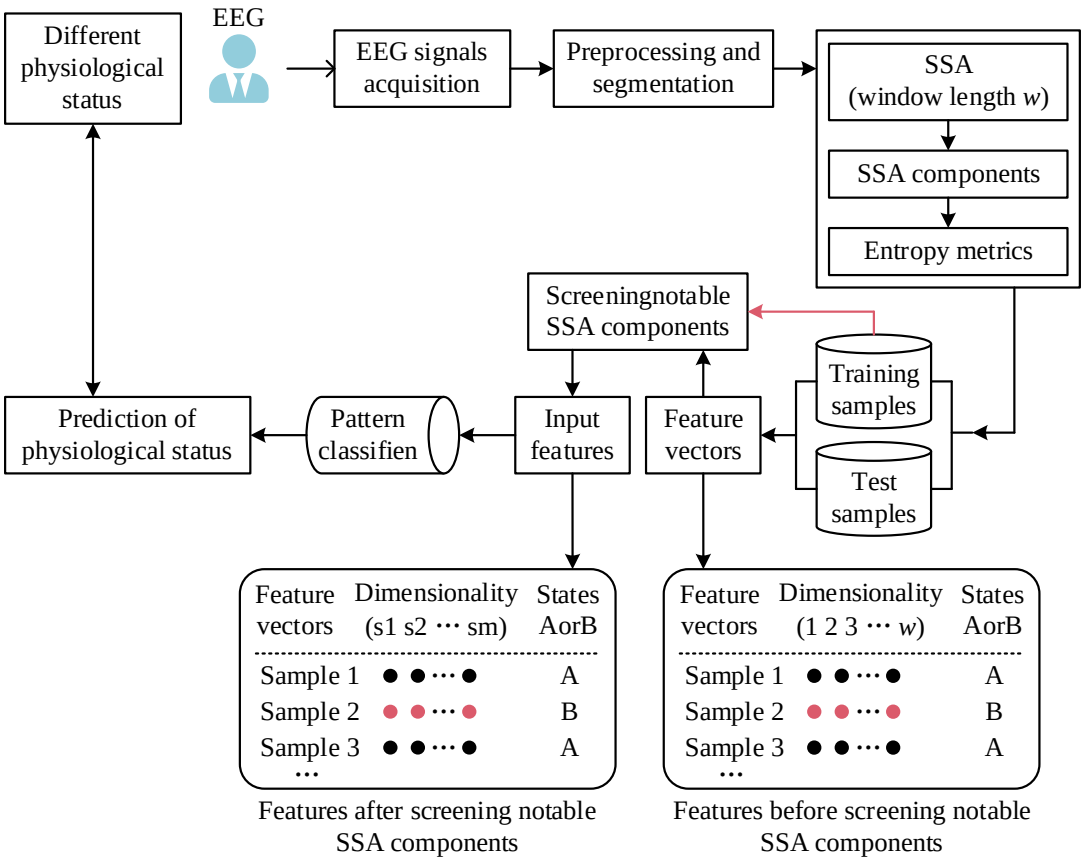


Figure 1: Framework for decoding electroencephalogram (EEG) signals

3.2 Singular Spectrum Analysis

The Singular Spectrum Analysis (SSA) method starts from the dynamic reconstruction of time series and constructs a trajectory matrix based on the fundamental idea of phase space

reconstruction. By performing singular value decomposition (SVD) on the trajectory matrix, the method achieves the goal of grouping feature components. Finally, the diagonal averaging method is used to reconstruct the feature components of different components in the time series.

The SSA method for time series consists of two complementary stages: decomposition and reconstruction. The SSA decomposition stage comprises two steps: embedding and SVD. In the embedding step of SSA decomposition, the phase space reconstruction method is primarily used to map the one-dimensional time series signal to obtain the trajectory matrix. Considering the time series $t(n), n = 1, 2, 3, \cdots, N$, the first step of the embedding process is to use phase space reconstruction with a window length of w to map the one-dimensional time series $t(n)$ to a multidimensional vector space, obtaining the trajectory matrix X of the time series $t(n)$, as follows:

$$X = \begin{bmatrix} t(1) & t(2) & t(3) & \cdots & t(L) \\ t(2) & t(3) & t(4) & \cdots & t(L+1) \\ t(3) & t(4) & t(5) & \cdots & t(L+2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t(w) & t(w+1) & t(w+2) & \cdots & t(N) \end{bmatrix} \tag{1}$$

In the SVD step of SSA decomposition, the trajectory matrix X is decomposed by singular value decomposition into the sum of w elementary matrices, that is:

$$X_j = \sqrt{\lambda_j} U_j V_j^T \tag{2}$$

$$X = X_1 + X_2 + \cdots + X_w = \sum_{j=1}^{w} \sqrt{\lambda_j} U_j V_j^T \tag{3}$$

where λ_j and U_j correspond to the eigenvalues and eigenvectors of the covariance matrix XX^T, respectively, and V_j is defined as $V_j = X^T U_j / \sqrt{\lambda_j}$. The vectors U_j and $V_j (j = 1, 2, ..., w)$ are called left and right singular vectors, respectively. The combination $(\sqrt{\lambda_j}, U_j, V_j)$ is defined as the triple of the trace matrix X [26]. It is worth noting that the eigenvalues λ_j are arranged in descending order, i.e., $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_w \geq 0$.

The SSA reconstruction phase, consists of two steps: grouping and diagonal averaging. In the grouping step, based on the singular value decomposition results, the singular values and their corresponding left and right singular vectors are grouped and combined, and the grouping matrix X_{c_k} can be obtained by referring to Equation (2). For the grouping of SSA reconstruction, each singular value with different sizes and its corresponding left and right singular vectors reflect different information of the time series and represent components with independent information. The specific operation process of the grouping step, based on the results of the singular value decomposition, forms a set of subscripts $\{1, 2, \cdots, w\}$ from the singular values of each order order $1, 2, \cdots, w$. Then, subset partitioning is implemented on the set of subscripts to obtain v subsets $c_1, c_2, \cdots, c_v (v \leq w)$. Finally, based on the mathematical definitions of Eqs. (2) and (3), the grouping matrix X_{c_k} is computed and combined into a new result matrix $\tilde{X}$.

$$\tilde{X} = \sum_{k=1}^{v} X_{c_k} \tag{4}$$

Since the dimension of the resultant matrix $\tilde{X}$ is $w \times L$, the resultant matrix $\tilde{X}$ can be rewritten according to the matrix elements in the form $\tilde{X} = (\tilde{x})_{i,j}$ ($1 \le i \le w$ and $1 \le j \le L$), that is:

$$\tilde{X} = (\tilde{x}_{i,j})_{i,j}^{w,L} = \begin{bmatrix} \tilde{x}_{1,1} & \tilde{x}_{1,2} & \tilde{x}_{1,3} & \cdots & \tilde{x}_{1,L} \\ \tilde{x}_{2,1} & \tilde{x}_{2,2} & \tilde{x}_{2,3} & \cdots & \tilde{x}_{2,L} \\ \tilde{x}_{3,1} & \tilde{x}_{3,2} & \tilde{x}_{3,3} & \cdots & \tilde{x}_{3,L} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{x}_{w,1} & \tilde{x}_{w,2} & \tilde{x}_{w,3} & \cdots & \tilde{x}_{w,L} \end{bmatrix} \tag{5}$$

The diagonal averaging step in the reconstruction phase of SSA is to average all diagonal elements $(\tilde{x})_{i,j}$ on the diagonal $i + j = n + 1$ of the result matrix $\tilde{x}$ to reconstruct to obtain a one-dimensional time series $y(n), n = 1, 2, 3, ..., N$, whose specific mathematical defining equation for $y(n)$, is as follows:

$$y(n) = \begin{cases} \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \tilde{x}_{i,n-i+1} & \text{for } 1 \le n < w \\ \dfrac{1}{w} \displaystyle\sum_{i=1}^{w} \tilde{x}_{i,n-i+1} & \text{for } w \le n \le L \\ \dfrac{1}{N-n+1} \displaystyle\sum_{i=n-L+1}^{N-L+1} \tilde{x}_{i,n-i+1} & \text{for } L+1 \le n \le N \end{cases} \tag{6}$$

The reconstructed time series signal $y(n)$ and the result matrix $\tilde{X}$ are one-to-one correspondence. In the grouping step of SSA reconstruction, if the condition $k = w$ is satisfied, the result matrix $\tilde{X}$ of Eq. (4) and the matrix X of Eq. (3) are equal. At this point, the original time series signal, $t(n)$, can be completely reconstructed from the resultant matrix $\tilde{X}$, i.e., $y(n) = t(n)$. More specifically, the subscript c_k of matrix X_{c_k} in Eq. (4) takes the value $1, 2, \cdots, w$ of the resultant matrix $\tilde{X}$ in turn will be defined as X_{c_k}, $c_k = 1, 2, ..., w$, respectively, and then, according to Eq. (6), the corresponding reconstructed signals $\tilde{y}_{c_k}(\eta)$, $c_k = 1, 2, \cdots, w$ are computed. under this condition, the time series signal $t(n)$ can be directly expressed as the sum of the one-dimensional time series $\tilde{y}_{c_k}(n)$ of each order of the reconstruction, viz:

$$t(n) = \sum_{c_k=1}^{w} \sum_{n=1}^{N} \tilde{y}_{c_k}(n) \tag{7}$$

3.3 SSA decomposition and reconstruction of EEG signals

In this paper, we investigate the integration of the SSA analysis step of EEG signals in the process of brain spot signal metric feature extraction. Consider the EEG time series signal $s(n), n=1,2,3,...,N$, the trajectory matrix of the EEG time series is obtained using phase space reconstruction with window length w. Using the eigenvalues of the SVD decomposition of the trajectory matrix and its corresponding left and right singular vectors, the corresponding SSA components of each order of the EEG signal $\tilde{s}_p(\eta), p=1,2,3,\cdots,w$ are reconstructed. Finally, by SSA decomposition and reconstruction, the EEG signal $s(n)$ can be represented in the form of the sum of w SSA components.

$$s(n)=\sum_{p=1}^{w}\sum_{n=1}^{N}\tilde{s}_p(n) \tag{8}$$

3.4 EEG entropy feature extraction map based on SSA method and entropy measure

The block diagram of EEG entropy feature extraction based on SSA components of EEG signal is shown in Fig. 2. In the EEG entropy feature extraction block diagram, the EEG signal is processed by the SSA method with four-step process of phase space reconstruction, SVD, grouping, and diagonal averaging, and then the SSA components of each order are obtained.The SSA components are all one-dimensional time sequences, and the entropy measure calculation can be directly carried out on the SSA components to realize the EEG entropy feature extraction. The experimental study in this chapter utilizes four entropy measurement algorithms for EEG entropy measurement calculation, which are approximate entropy, sample entropy, fuzzy entropy and multi-scale entropy.

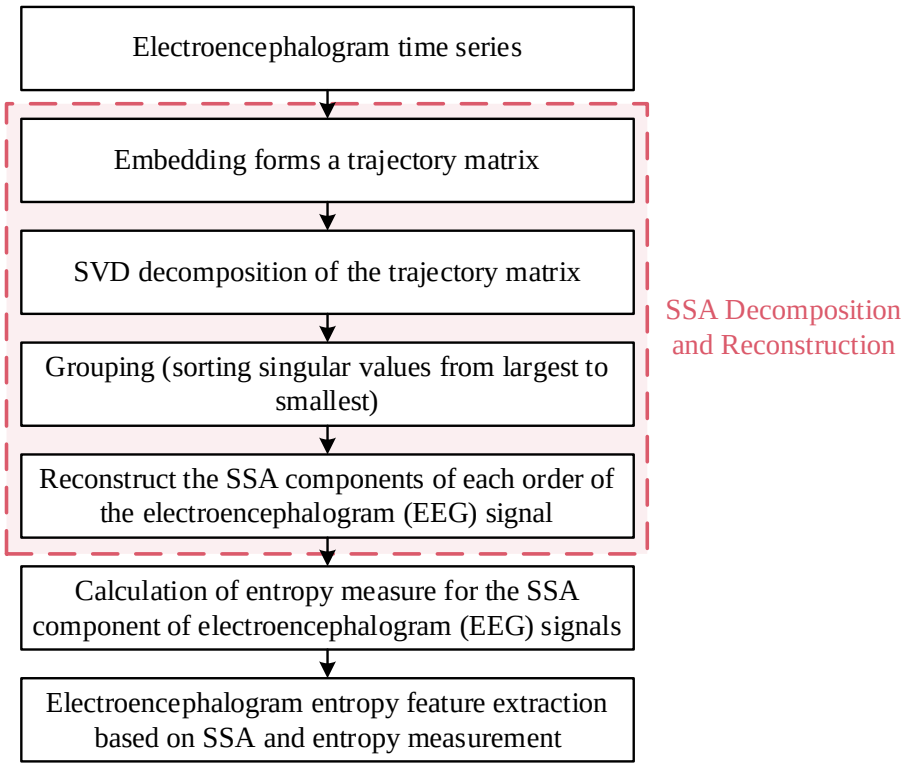


Figure 2: Framework of EEG entropy features extraction based on SSA and entropy measures

4 Topological constraint retention based method for CSP emotional EEG prediction

4.1 EEG Spatial Characterization Method Based on Common Spatial Pattern (CSP)

(1) Traditional CSP feature extraction technique

CSP algorithm as an efficient spatial filtering algorithm, its mathematical basis is the simultaneous diagonalization of the two matrices, and find the optimal spatial projection to maximize the difference between the two types of signals. Let the EEG data D, D_1 and D_2 be two types of multi-channel EEG data samples respectively, and their dimensions are $M \times T$, M is the number of electrodes, i.e., the number of channels, and T is the number of samples collected for each channel.The procedure of CSP algorithm is as follows:

(a) Calculate the covariance matrices D_1 and D_2 normalized by R_1 and R_2 as follows:

$$R_1 = \frac{D_1 D_1^T}{trace(D_1 D_1^T)}, R_2 = \frac{D_2 D_2^T}{trace(D_2 D_2^T)} \tag{9}$$

where D^T denotes the transpose matrix of matrix D and $trace(D)$ denotes the sum of elements on the diagonal, i.e., the trace of matrix D. Using $\overline{R_1}$ to denote the average covariance matrix for the first class of sample data, $\overline{R_2}$ to denote the average covariance matrix for the second class of samples, and R_c to denote the mixed spatial covariance matrix for the two classes of data, we have:

$$R_c = \overline{R_1} + \overline{R_2} \tag{10}$$

(b) Eigenvalue decomposition of the mixed-space covariance matrix R_c:

$$R_c = U \Lambda U^T \tag{11}$$

where U denotes the eigenvector matrix, Λ is the corresponding diagonal array formed by the eigenvalues. After the eigenvalues are arranged in descending order and the whitening transformation U is performed, the matrix of whitening values can be obtained:

$$P = \sqrt{\Lambda^{-1}} U^T \tag{12}$$

Acting the whitened value matrix P on R_1 and R_2 yields:

$$S_1 = P R_1 P^T, S_2 = P R_2 P^T \tag{13}$$

S_1, S_2 have common eigenvectors, and then a principal component decomposition of S_1 and S_2 is obtained:

$$S_1 = B \Lambda_1 B^T, S_2 = B \Lambda_2 B^T \tag{14}$$

where B is the eigenvector matrix and the sum of the diagonal array of the two eigenvalues Λ_1 and Λ_2 is the unit matrix, i.e., satisfies the following condition

$$\Lambda_1 + \Lambda_2 = I \tag{15}$$

where I is the unit matrix.

(c) Calculate the projection matrix: for the eigenvector matrix B, when a category S_1 has the largest eigenvalue, the corresponding eigenvector S_2 has the smallest eigenvalue, so the projection matrix can be obtained by using the matrix B to realize the classification of the two categories of problems:

$$W = (B^T P)^T \tag{16}$$

(d) Calculate the feature matrix: a raw EEG data $D_{M\times T}$ can thus be obtained by projecting it through the projection matrix W:

$$Z_{M\times T} = W_{M\times M} D_{M\times T} \tag{17}$$

(2) CNN-based CSP characterization methods

(a) Principle of convolutional neural network (CNN) technology

The convolutional layer in a CNN is for feature extraction from the input data and contains multiple convolutional kernels within it. A convolutional kernel is similar to a neuron in a feed-forward neural network, and each element that makes up the convolutional kernel corresponds to a weight coefficient and a deviation. Each convolution kernel is convolved with the upper layer of data, and after the activation function, the bias value is added to form a local representation of the feature map in this layer. The process of its convolution operation can be expressed by the following equation:

$$M_j^L = f\left(\sum_{i\in N^{L-1}} M_i^{L-1} * C_j^L + b_j^L\right) \tag{18}$$

where * denotes the convolution operation, b denotes the bias value, and M_i^{L-1} and M_j^L denote the inputs and outputs of the L th layer, also known as the feature map. C_j^L is the j th convolutional kernel in the L th convolutional layer. $f(x)$ is the activation function, typically a sigmoid function or hyperbolic tangent function, etc. Its objective function is the average value of the loss function in the whole dataset D.

$$Loss = \frac{1}{|D|}\sum_{i}^{|D|} f_L(X^{(i)}) \tag{19}$$

where Loss is the loss function, according to the last weight value W_f, the stochastic gradient descent method is used to optimize W with the following formula:

$$W_{t+1} = W_t - \alpha Loss(W_t) \tag{20}$$

where α is the learning rate.

(b) CNN-based feature optimization method

In this section, a CNN-based feature optimization method is proposed, in which the feature matrix is first learned through a CNN network, and the effective features in the feature matrix are screened by the distribution of the weights of the fully connected layer in the following steps:

Assume that the input feature graph of the fully connected layer in CNN is s with dimension $n \times l$ and the t th fully connected layer with k_t nodes, which is denoted by w_t. Then when $t = 1$, there is only 1 fully connected layer w_1 for the beneficial weighted weight matrix.

When $t > 1$, the lower layer weights are used as weights to weight the upper layer weights The beneficial weighting weight matrix is:

$$W = W_1 \times W_2 \times \cdots \times W_t \tag{21}$$

where the h th column denotes the weighted weights useful for categorizing as h. Rearranging them yields $W_h^{'}$:

$$[s \times n \times l, 1] \to [s, n, l] \tag{22}$$

At this point, the matrix $W_h^{'}$ can be viewed as the s amplitude $n \times l$ weights image, obtained by taking the absolute values:

$$W_h^{''}(t,i,j) = \left| W_h^{'}(t,i,j) - \frac{1}{n*l} \sum_{i=1}^{n} \sum_{j=1}^{l} W_h^{'}(t,i,j) \right| \tag{23}$$

The degree of deviation from the mean is obtained, highlighting the effective weights. Then the average of all images is obtained:

$$T_h(i,j) = \frac{1}{s} \sum_{t=1}^{s} W_h^{''}(t,i,j) \tag{24}$$

For the matrix T_h, find the standard deviation of each row, with i taking the range $(1 \le i \le n)$:

$$\overline{T_h}(i) = \frac{1}{l} \sum_{j=1}^{l} T(i,j) \tag{25}$$

$\overline{T_h}$ is a column vector that represents the contribution of each class of features of the feature matrix to the classification, and the larger the contribution indicates that the feature is more effective for classification h.

Therefore, the degree of validity of the i th feature of the feature matrix for all classifications can be obtained $\overline{X}(i)$:

$$\bar{X}(i)=\frac{\sum_{j_1-1}^{t_k-1}\sum_{j_2-j_1+1}^{t_k}|\overline{T_{j_1}}-\overline{T_{j_2}}|}{\frac{1}{2}t_k^2-\frac{1}{2}t_k} \tag{26}$$

In order to find the most efficient feature, the deviation is solved for all rows to obtain the column vector P.

$$P(i)=\left|\frac{\bar{X}(i)-\frac{1}{n}\sum_{i=1}^{n}\bar{X}(i)}{\frac{1}{n}\sum_{i=1}^{n}\bar{X}(i)}\right| \tag{27}$$

Sort P and the rows that satisfy $P(i)\geq\varphi$ are the ones sought (φ is the threshold, usually the standard value of deviation of 0.5 is considered to be a valid feature, so $\varphi=0.5$ is taken).

4.2 TCM-CSP Emotional EEG Calculation Method Based on Topological Constraint Holding

(1) Definition of the problem

Suppose that EEG data were collected from S subjects, each subject $S_k, k=1,\cdots,S$ has p segments (epochs) of EEG with signals as $\{E_1,E_2,...,E_p\}$. Assuming that a total of m electrodes are included, each segment contains consecutive sampling points as $E_a=\{D_a^b, b\in[1,m]\}$, where D_a^b denotes the a th segment of EEG data in the b th electrode. Meanwhile, the EEG signal of a single electrode b can be represented as $\xi(i)$, where i denotes a sample of the time series, i.e., $i\in[1,n]$.

Therefore, $\xi(i)$ can also be expressed as $\{D_a^b, a\in[1,p]\}$. Assuming that there are a total of A affective states, the labeling of m segments can be defined as $\{L_c^a\}_{a=1}^p, c\in A$, where L_c^a denotes that the labeling of the a th segment belongs to the c th affective state, and A denotes the set of all the affective states induced by auditory signals. Assuming that the affective states of all p-segment EEGs are in the set A, then the affective states of all segments in order can be formed into the set Z, which is denoted as $\{L^a\}_{a\in Z}$, and the p-segment EEG signals $\{E_a\}_a=1^p$ affective The state identification problem can be represented as the following mapping:

$$F:E\rightarrow L \tag{28}$$

(2) Emotion correlation analysis

To characterize emotion-related EEG, the correlation between EEG responses and emotion categories was first calculated.Spearman rank correlation coefficients were used between EEG features (BPC) and emotion labels (L). For samples of size n, the n raw scores BPC, L were converted to ranks bpc_i, l_i and ρ were calculated from the following equation:

$$\rho = \frac{\sum_i (bpc_i - \overline{bpc})(l_i - \overline{l})}{\sqrt{\sum_i (bpc_i - \overline{bpc})^2 \sum_i (l_i - \overline{l})^2}} \tag{29}$$

where BPC denotes the feature vector of EEG signal and L denotes the emotion label. The higher the value of ρ, the greater the degree of correlation. Based on this feature the most relevant feature set can be found.

$$\hat{BPC} = \{bpc \,||\, \rho \,|> Th_\rho\} \tag{30}$$

where Th_ρ is the threshold of ρ.

(3) Definition of topological constraint matrix TCM

The basic idea of constraint matrix is that a feature constraint matrix is established first, where the elements represent the relevance of the dimensional feature to the emotion, respectively, so that the feature results can be constrained during feature extraction.

For the feature vector matrix B, where $B = B_0, B_1, \cdots, B_{N-1}$, the constraint matrix $P = p_{ij}$ is an $n \times m$ matrix.

$$p_{ij} = \begin{cases} 1, p_{ij} \geq Th_\rho \\ 0, p_{ij} \leq Th_\rho \end{cases} \tag{31}$$

From the definition, the constraint matrix is a Boolean matrix that expresses the constraints on the correlation of the features with the emotional response. For the element p_{ij} in the matrix, if the correlation between the corresponding feature and the emotional response in the i th row is greater than Th_ρ, then the value of this element is 1. Otherwise, the value of this element is 0. According to the constraint rule, this constraint matrix can be generated, and the principle of the topological constraint matrix computation is shown in Figure 3.

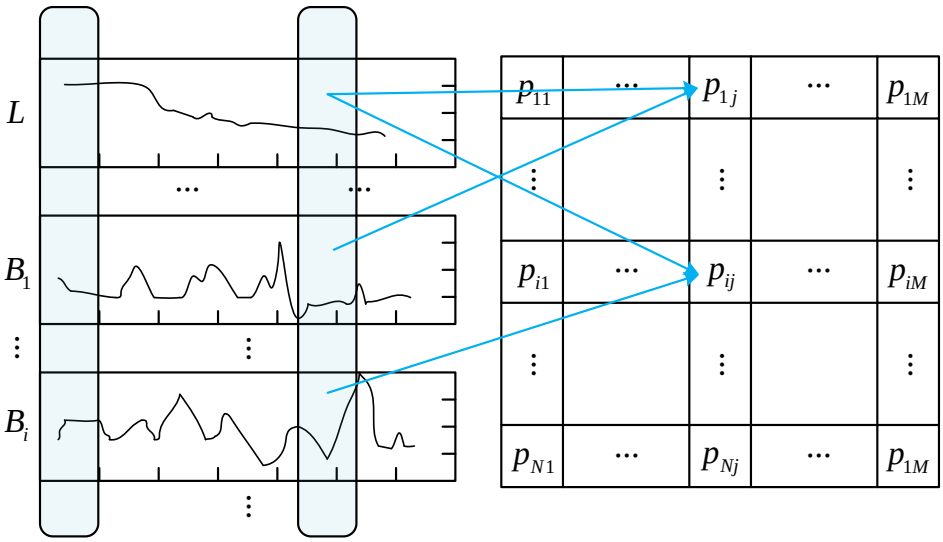


Figure 3: The principle of topology constraint matrix calculation

(4) TCM-CSP feature extraction method

The EEG features are computed with a constraint matrix before CSP projection, and the features are optimized by matrix masking operation, the schematic diagram of TCM optimized CSP is shown in Fig. 4.

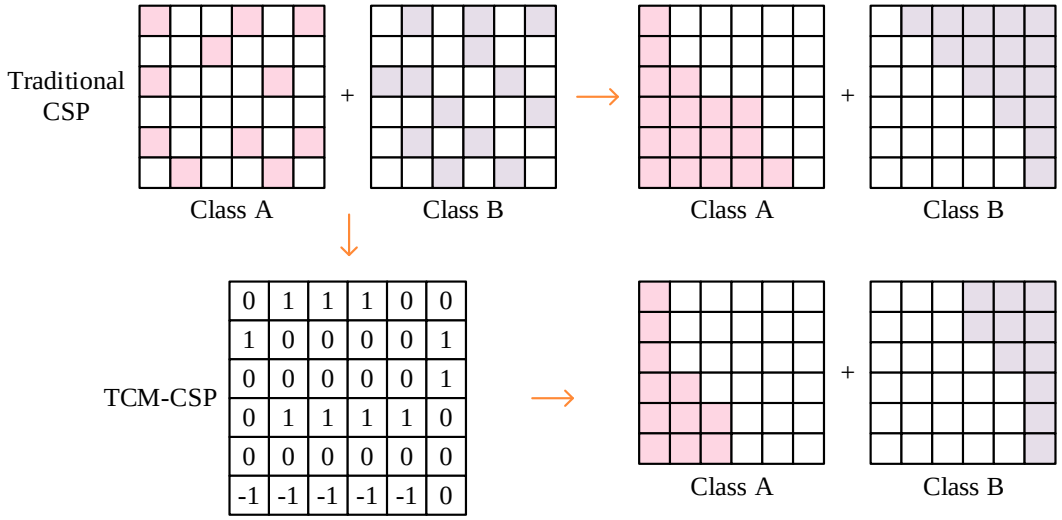


Figure 4: Schematic diagram of TCM optimizing CSP

For the EEG analysis problem studied in this paper, let the EEG data can be represented as a matrix $E_{M\times T}$, with M representing the number of electrodes and T representing the number of sampling points, and then the masking operation of the constraint matrix can be expressed as:

$$E' = E_{M\times T} * P \tag{32}$$

The recalculated spatial covariance matrix can be expressed as:

$$R = \frac{E'E'^{T}}{trace(E'E'^{T})} \tag{33}$$

For the recognition classification problem, the covariance R can be obtained by averaging the results of multiple experiments, and the spatial covariance of the two classes of datasets can be expressed as $R_c = \overline{R_1} + \overline{R_2}$. where R_c denotes the sum of the spatial covariance of the two types of sample data, $\overline{R_1}$ denotes the spatial covariance of the data of the first type of sample, and $\overline{R_2}$ denotes the spatial covariance of the second type of sample.

Let $R_c = U_p \Lambda_c U_c^T$, the eigenvector matrix can be expressed as U_c and the diagonal array of eigenvalues is Λ_c. The whitening transformation U_c can be obtained:

$$P = \frac{1}{\sqrt{\Lambda_c}} U_c^T \tag{34}$$

5 Electroencephalographic data processing

EEG signals are non-smooth random signals and are susceptible to various noise interferences. So better acquisition of swimming information reflecting brain activity and state is an important prerequisite for EEG analysis. Firstly, the EEG signal was preprocessed using the EEGLAB toolbox of MATLAB. The sampling rate was set to 1000 Hz, and the preprocessing operations such as filtering, artifact removal, Run ICA, and eye removal were performed in EEGLAB.

5.1 Filtering

Since the currently recognized EEG frequencies are mainly concentrated below 50Hz, 1Hz to 50Hz bandpass filtering is required to remove DC interference and high frequency signals. Notch filtering is performed before filtering for removing 55Hz industrial frequency interference. This part of the 55Hz signal comes from the interference of the utility 55Hz AC power to the equipment. So a band pass filter of [1, 40] is required.

5.2 Artifact removal

The entire waveform is examined and artifacts are removed manually by hand. Some obvious artifactual signals can be recognized by the human eye. Electromyographic (EMG) signals are mainly interference signals generated by head, limb, jaw, tongue, and swallowing movements.EMG signals often have a frequency greater than 30 Hz, and show a wide separation of conductances and many burr signals in the longitudinal direction. EMG signals in the head and face are predominantly in the frontal or temporal regions. Manual removal of the artifacts allowed for cleaner data before RUN ICA and easy access to the electrooculographic component. The figure below shows a comparison of the EEG before and after artifact removal.

5.3 Run ICA and remove independent source noise components

ICA is a method used to find hidden factors or components from multivariate statistics and is considered an extension of principal component analysis and factor analysis. The blink component is shown in Figure 5. The eye drift component is shown in Figure 6. The head movement component is shown in Figure 7.

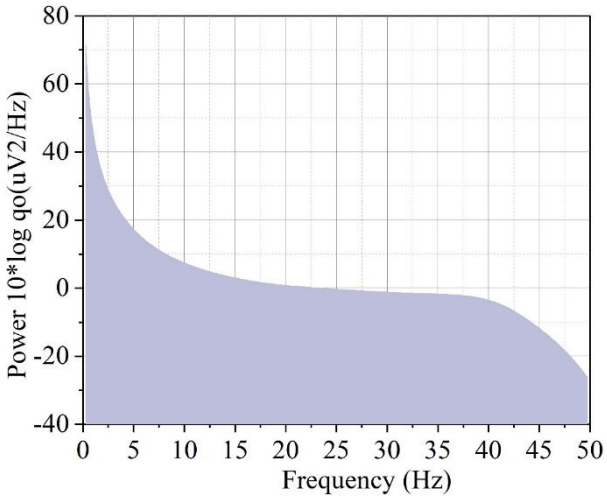


Figure 5: Blinking ingredient

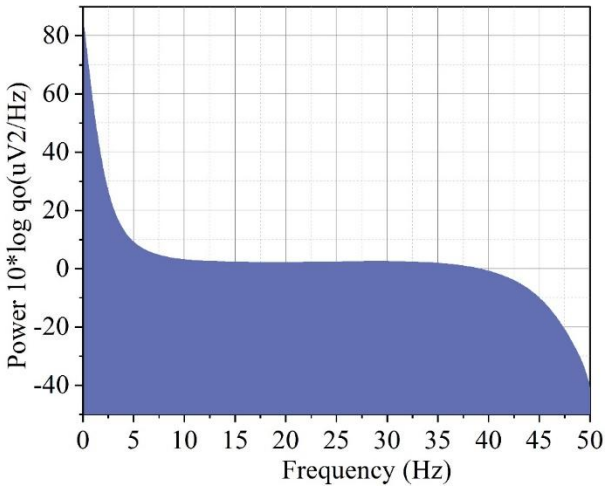


Figure 6: Eye drift

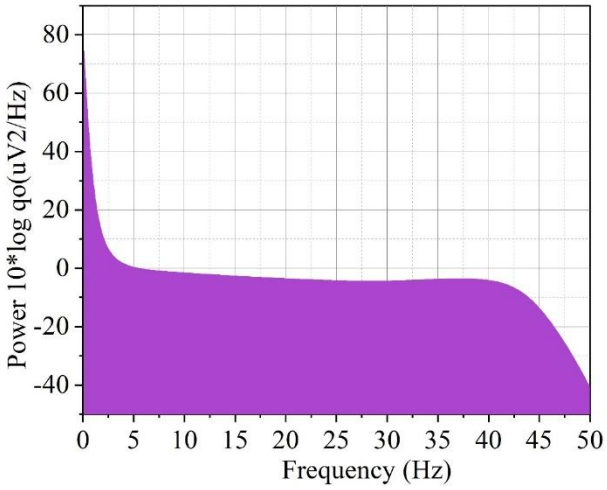


Figure 7: Head component

5.4 Analysis of the results of EEG experiments

Taking red music as an example, the processing results of sample entropy showed that, under red audiovisual synchronized stimulation, the lead sample entropy was slightly higher than that of visual stimulation alone, except for the lead sample entropy of frontal pole FP1 and FP2, central parietal lobe CP1 and CP6. The average value of audiovisual sample entropy of most other brain regions was low with red video alone stimulation. The lowest sample entropy value for music indicates that the brain activity under a single auditory stimulus has the lowest time series complexity and is more regular compared to the visual stimulus and the audiovisual year-on-year stimulus. The red sample entropy mean is shown in Figure 8.

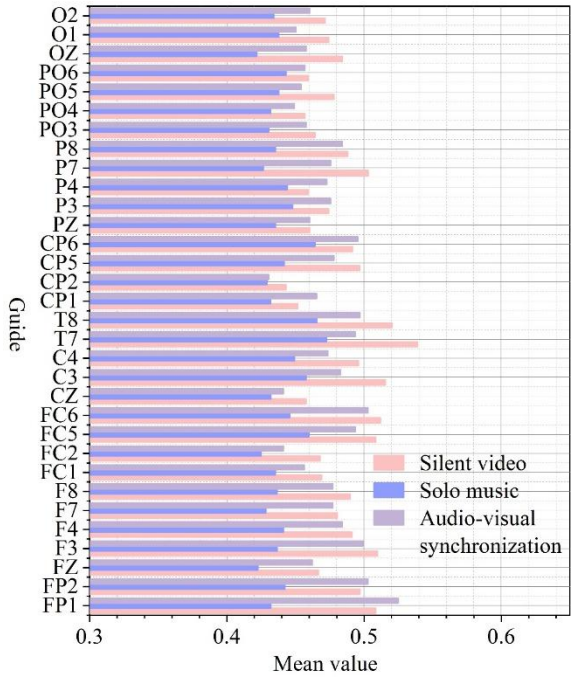


Figure 8: The average entropy of the red sample

The three-color sample entropy averages are shown in Figure 9. The red audiovisual stimulus sample entropy values were lower than the blue and green except in the FC2 and C3 leads. The red audiovisual stimulus sample entropy values in most of the leads are slightly higher than the purple and blue results. And the difference between the values of purple and blue is not large, indicating that the musical visualization performance of purple and blue audiovisual stimuli is more regular compared to red, which has a certain effect on the soothing regulation of brain emotion.

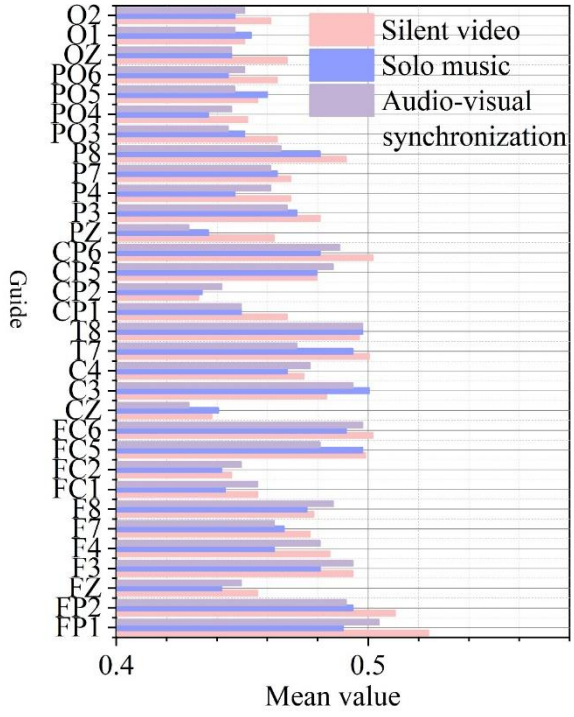


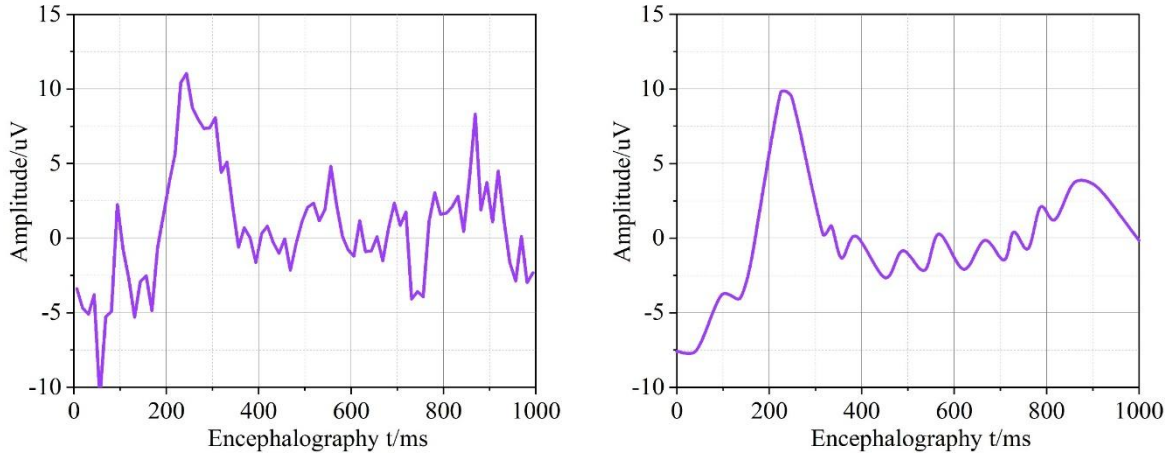
Figure 9: The average entropy of the three-color samples

6 Synchronization analysis of auditory associations and musical emotions

Studies have shown that the prefrontal lobe is a functional brain region associated with emotions, so during the analysis of EEG synchronization using the methods in Chapter 4, the phase synchronization between the Fp1 and Fp2 leads located in the prefrontal lobe and the other leads was mainly analyzed. The following are the results and analyses of EEG feature extraction and visual, auditory, and audio-visual synchronization-induced EEG synchronization, respectively.

6.1 EEG feature extraction results

The waveforms of occipital area lead EEG after coherent averaging and wavelet reconstruction are shown in Fig. 10 (Fig. a shows the EEG waveform after coherent averaging processing, and Fig. b shows the waveform of α rhythm after wavelet reconstruction). Figure (a) shows the evoked EEG signals obtained after modeling and 20 times of coherent averaging of the EEG data in the occipital lead in the pleasant state, and Figure (b) shows the waveforms of EEG features reconstructed by wavelet transform. As can be seen from the figure, after wavelet transform, the feature signals become smoother and more distinctive, which can effectively reduce the pseudo-synchronization phenomenon caused by burrs. Moreover, the event-related potential signal feature appeared at 285ms, indicating that the brain integrated the information of the received audiovisual stimulus 285ms after the stimulus appeared.



(a) The brain waves that are treated with the same average

(b) The α rhythm waveform after the wavelet reconstruction

Figure 10: The waveform after coherent averaging and wavelet reconstruction

6.2 Significant difference analysis

In this paper, paired-sample t-test analysis was conducted to analyze the EEG phase synchrony indices during different emotional cognitions. The results show that there is a significant difference between the EEG synchronization indexes of pleasant emotion cognition and anger, sadness, surprise, disgust, and fear emotion cognition (negative emotion). t-test results of the EEG phase synchronization indexes of emotion cognition during pleasant and fearful emoticon pictures are shown in Table 1, from which it can be seen that there is a significant difference between the EEG synchronization of these two kinds of emotion cognition (positive emotion

and negative emotion) (Sig. < 0.05). While there is no significant difference between the synchronization indices of anger, sadness, surprise, disgust, and fear during the cognition of these five negative emotions, the t-test results of the EEG synchronization indices of the emotion cognition during the disgust and fear emoticon pictures are shown in Table 2, from which it can be seen that there is no significant difference between the EEG synchronization of the cognition of these two emotions (Sig. > 0.05). Therefore in this paper only the synchronization of positive and negative emotional cognition is analyzed.

Table 1: The results of the electrical phase synchronization index t test

Lead link	Mean	Standard deviation	The standard error of the mean	t	Sig.(Double side)
Fp1	0.1129	0.1431	0.0299	2.189	0.074
Fp2	-0.4078	0.3042	0.1099	-4.092	-0.003
F3	-0.0237	0.2641	0.0685	-0.062	0.955
F4	-0.2953	0.283	0.0833	-3.298	0.014
T3	-0.3682	0.4627	0.1504	-2.427	0.046
T4	-0.2435	0.3053	0.0978	-2.809	0.021
P3	-0.236	0.2868	0.0749	-2.595	0.028
P4	-0.3469	0.2732	0.0856	-4.217	0.021
O1	-0.264	0.2923	0.093	-2.681	0.027
O2	-0.235	0.3011	0.1062	-2.519	0.005

Table 2: The results of the brain electrosynchronous index t test of emotional cognition

Lead link	Mean	Standard deviation	The standard error of the mean	t	Sig.(Double side)
Fp1	-0.1722	0.2263	0.0694	-2.276	0.044
Fp2	-0.2101	0.2192	0.0732	-3.393	0.001
F3	0.0277	0.2297	0.0628	0.486	0.637
F4	0.1475	0.268	0.0955	1.709	0.136
T3	-0.2012	0.3068	0.094	-1.705	0.127
T4	-0.1115	0.4463	0.1381	-0.774	0.476
P3	-0.1375	0.3175	0.0873	-1.421	0.175
P4	-0.1696	0.2278	0.0767	-2.374	0.05
O1	-0.137	0.3392	0.1147	-1.183	0.279
O2	0.1827	0.2876	0.0954	2.321	0.039

6.3 Analysis of visually evoked EEG synchronization

The EEG phase synchronization indices for pleasant emotion cognition and aversive emotion cognition are shown in Table 3. It can be seen that the Fp1 lead showed high phase synchronization (γ_H >0.6) with the P3, P4, and O1 leads, while the Fp2 lead showed high phase synchronization (γ_H >0.6) with the O2 lead, indicating that there was a high degree of synchronization between the left frontal lobe and parietal-occipital lobe, and between the right frontal lobe and the occipital lobe during the completion of the cognition of pleasant emotion with a single visual stimulus, and that the right frontal lobe and the occipital lobe had a high degree of synchronization between the left frontal lobe and parietal-occipital lobe. Fp2 leads showed high phase synchronization with F4, P3, and O2 leads (γ_H >0.65), while Fp1 leads did

not show high synchronization with these leads (γ_H <0.5), suggesting that only the right frontal lobe of the brain and the parieto-occipital lobe were highly synchronized with each other when cognition of aversive emotions was accomplished with a single visual stimulus.

Table 3: Electroencephalogram phase synchronization index

		Fp1	F3	F4	T3	T4	P3	P4	O1	O2
Cognition of pleasant emotions	Fp1	0.193	0.501	0.136	0.212	0.541	0.658	0.778	0.620	0.064
	Fp2	0.191	0.398	0.355	0.489	0.322	0.301	0.355	0.396	0.634
Aversion emotion cognition	Fp1	0.420	0.563	0.501	0.327	0.389	0.424	0.260	0.145	0.386
	Fp2	0.420	0.508	0.684	0.451	0.367	0.894	0.525	0.484	0.890

According to the study of the Brodmann partitioning system of the brain and the location and function of each area, it is known that the prefrontal lobe of the brain is related to emotions, the occipital lobe is related to visual stimulation, the temporal lobe is related to auditory stimulation, and the parietal lobe is the sensory area that synthesizes all kinds of sensations. Combined with the above experimental analysis, it can be seen that (1) when receiving the stimulation of pleasant expression pictures, the visual area and sensory area of the brain and the left emotional area produced a synchronous oscillation phenomenon, from which it can be inferred that the cognition of pleasant emotions is mainly related to the left frontal lobe. While the right emotional area had a higher synchronization with the visual area, it was presumed that the synchronization was produced by the simultaneous reception of visual stimuli, and this process did not involve the sensory area. (2) When receiving the stimulation of disgust expression pictures, the visual area, sensory area and right emotional area of the brain produced a synchronous oscillation phenomenon, from which it can be inferred that the cognition of disgust emotion is mainly related to the right frontal lobe. And the synchronization between the left emotional region and all other brain regions was not high, indicating that the left frontal lobe was not involved in the cognitive task of disgust emotion.

6.4 Analysis of auditory-evoked EEG synchronization

The EEG phase synchronization indices for pleasant emotion cognition and fearful emotion cognition are shown in Table 4. As can be seen from the comparison plots, the Fp1 leads showed higher phase synchronization with the F3, T3, T4, and P4 leads showed higher phase synchronization (γ_H >0.6), while the Fp2 leads showed higher phase synchronization with the F4, T3, and T4 (γ_H >0.6), indicating that the left frontal lobe of the brain and the parietal-temporal lobe had a high degree of synchronization of electrical activity and the right frontal lobe and temporal lobe also had a high degree of synchronization of electrical activity when a single auditory stimulus was given to complete the cognition of the pleasant emotion. The Fp2 leads showed high phase synchrony (γ_H >0.65) with the F4, T3, and T4 leads, while the Fp1 leads did not show high synchrony (γ_H <0.50) with these leads, suggesting that only the right frontal and parietal-temporal lobes of the brain showed high synchronized electrical activity when a single auditory stimulus was given to complete the cognition of fearful emotions. This suggests that only the right frontal and parietal temporal lobes of the brain have highly synchronized electrical activity when a single auditory stimulus is given to complete the cognition of fear. Combined with the above experimental analyses, it can be seen that: (1) when receiving pleasant sound stimuli, the auditory and sensory regions of the brain synchronized

with the left emotional region, and it can be inferred that the cognition of pleasant emotions is mainly related to the left frontal lobe. The higher synchronization between the right emotional region and the auditory region may be due to the synchronization generated by the auditory stimulus without the participation of the sensory region. (2) When the fearful sound stimulus was received, the auditory area, the integrated sensory area and the right emotional area oscillated synchronously, and it can be inferred that the cognition of the fearful emotion is mainly related to the right frontal lobe. The synchronization between the left emotional region and other brain regions was not high, indicating that the left frontal lobe was not involved in the cognitive task of fearful emotion.

Table 4: Electroencephalogram phase synchronicity index

		Fp2	F3	F4	T3	T4	P3	P4	O1	O2
Cognition of pleasant emotions	Fp1	0.440	0.760	0.525	0.943	0.925	0.407	0.634	0.329	0.356
	Fp2	0.453	0.321	0.654	0.797	0.745	0.449	0.362	0.558	0.424
Aversion emotion cognition	Fp1	0.486	0.294	0.113	0.447	0.498	0.053	0.387	0.276	0.393
	Fp2	0.473	0.537	0.809	0.857	0.883	0.661	0.284	0.500	0.047

6.5 Analysis of audiovisually evoked EEG synchronization

The EEG phase synchronization indices during pleasant emotion perception and sad emotion perception are shown in Table 5.The Fp1 leads showed high phase synchronization with the F3, P3, P4, and O1 leads (γ_H >0.7), whereas the Fp2 leads did not show high synchronization with these leads (γ_H <0.4), suggesting that the left frontal and parieto-occipital lobes of the brain showed high synchronized electrical activity when receiving synchronized stimulation of the pleasant expression pictures and sounds. Fp2 leads showed high phase synchronization with T3, T4, P3, O1, and 02 leads (γ_H >0.6), while Fp1 leads did not show high synchronization with these leads (γ_H <0.4), indicating that the right frontal lobe of the brain and the parieto-occipital and subversive lobes of the brain showed high synchronization during the synchronous stimulation of pictures of sad expressions and sounds. Combined with the above experimental analysis, it can be seen that (1) when receiving pleasant expression pictures and sound synchronous stimulation, the parieto-occipital lobe and the left frontal lobe produced high EEG phase synchronization, i.e., the visual area and the left emotional area produced the phenomenon of synchronous oscillations, while the temporal lobe and the left frontal lobe did not have high synchronization, i.e., the phenomenon of synchronous oscillations was not obvious. From this, it can be inferred that the visual channel plays a dominant role during pleasant emotion cognition. (2) When receiving sad expression pictures and sound synchronous stimulation, the occipital lobe and temporal lobe both produced high EEG phase synchronization with the right frontal lobe, i.e., the visual area and the auditory area produced synchronous oscillation phenomenon with the right emotional area, but the synchronization between the temporal lobe and the right frontal lobe was higher compared to the synchronization between the occipital lobe and the right frontal lobe.

Table 5: Brain electrophase synchronicity index of happy emotional cognition and sadness

		Fp2	F3	F4	T3	T4	P3	P4	O1	O2
Cognition of pleasant emotions	Fp1	0.168	0.746	0.370	0.444	0.127	0.955	0.946	0.960	0.248
	Fp2	0.161	0.320	0.381	0.477	0.415	0.152	0.094	0.110	0.316
Aversion emotion cognition	Fp1	0.352	0.228	0.260	0.314	0.289	0.139	0.327	0.217	0.211
	Fp2	0.347	0.430	0.601	0.859	0.919	0.650	0.578	0.693	0.711

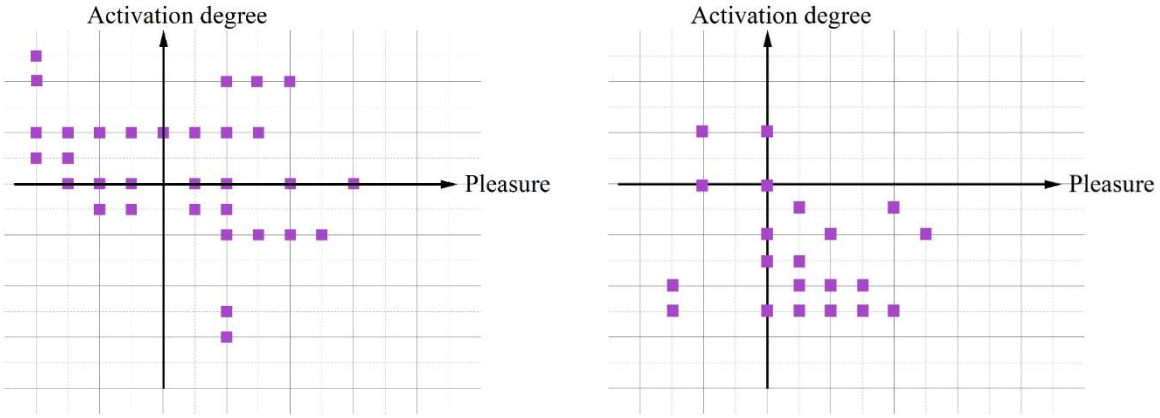
7 Auditory-associative musical-emotional experiments

7.1 Experimental sample construction

The construction of the samples needs to take into account the scenarios of the subjects and the acceptability of a short period of time, and it needs to be made easy to understand and convenient for the collection and generalization of the experimental data. In the preliminary stage of the experiment, it is necessary to determine the emotion model and assessment form used to collect data, set up the sound samples played in the experiment, the descriptors used in the assessment, and the picture samples viewed in the experiment. In order to test the emotional response brought about by the changes produced by the timbre of various types of musical instruments, four different types of instruments, namely, grand piano, oboe, guitar, and bells, were selected among keyboard, wind, plucked, and percussion instruments to perform the same sound, resulting in four basic sound segments. On the basis of these four sounds, we used the Fuzzy effect of the software, and adjusted the knobs of Fuzzy, Level, and Tone to the values in the "Fuzzy Overload" mode to create four distorted timbre effects. Adjust the reverb of the four base sounds to the maximum value to create a four-band reverb tone effect. The above sample settings create two groups of sounds with altered timbres.

7.2 Experimental Procedure and Analysis

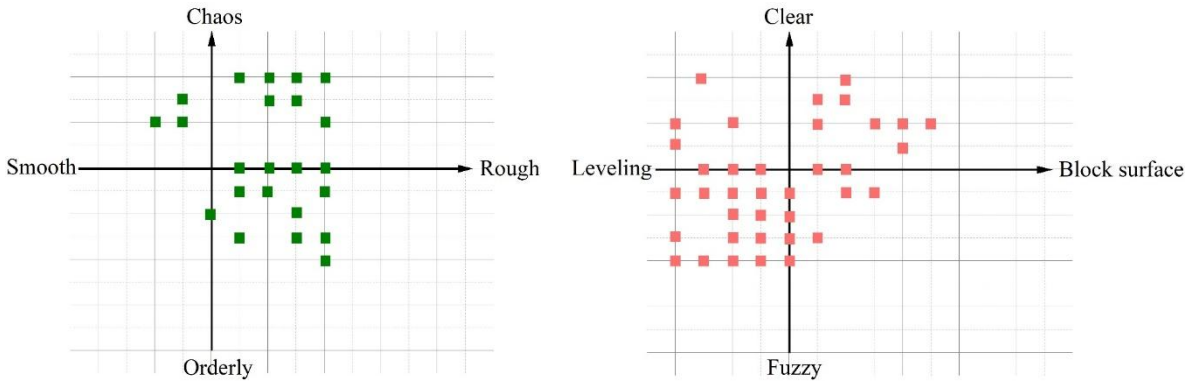
(1) Emotional tendency analysis: The emotional tendency of the distortion group and the reverb group is shown in Fig. 11 (Fig. a is the distortion group and Fig. b is the reverb group). From the distribution of dot plots after overlapping the data, the main emotional areas generated by the subjects' experience can be seen. It can be seen that after the overall change of timbre in the distortion group and the reverb group, there is a uniform change in the concentrated areas of emotion, and visually each group of sounds has a similar tendency. Under the reverb effector, the emotional experiences recorded by the subjects were mostly high pleasantness, low arousal states, and according to the cyclic affective model, the emotional experiences were mostly relaxing, tranquil, and serene. To summarize this item, the acoustic group showed differences in the basic emotions of all four instruments, which can be used as the emotional origin of the data for reference. The distortion group was more agitated (-V+A), and the reverberation group tended to be more tranquil and relaxed (+V-A), and the two groups showed a negative correlation between the two emotions, which can be seen in the great influence of the timbre of the different effect treatments on the emotional perceptions.



(a) Distortion Group (b) Reverberation group

Figure 11: Distortion group and reverberation group emotional trend

(2) Visual tendency analysis: the four quadrants delineated are a range of emotional tendencies, an emotional tendency; the stronger a certain timbre is expressed, the stronger the emotional arousal in this region, and at the same time, the more it tends to be polarized towards a certain dimension in the visual embodiment. In the corresponding visualization, the distortion group overall is hovering on the mediating line. The overall distribution of ordered-chaotic on the vertical axis is more even, but the number of samples selected for chaotic features is higher, and it can be assumed that the overall expression of the distorted timbres tends to be rough and on the chaotic side. In the reverb effector, the overall performance of the reverb timbre is flat and fuzzy tendency, in which the performance of sound four is relatively special, should be related to the qualities of the percussion instrument itself, the short collision sound of the object is more likely to produce the feeling of clarity and certainty, and in the visual expression of the instrument itself can be taken into account in the expression of the timbre of the instrument itself. The visual trend of the distortion group and the reverb group is shown in Figure 12 (Figure a is the distortion group and Figure b is the reverb group).



(a) Distortion Group (b) Reverberation group

Figure 12: Visual trends of the distortion group and the reverberation group

(3) Outline color impression: In the test of the acoustic group, each subject was asked to draw down the outline and color impression of this sample, which can form the preliminary observation and help the subsequent setting of the base image. The contour expression drawn

by the subjects was relatively abstract, with basically no figurative shapes. After summarizing the formal characteristics of the various types of images, four types were classified under the dimension of contour shapes, angular and closed, angular and closed, open and scattered. Sound I contour impressions were most heavily weighted as angular and closed. Sound two was angular and closed with the largest weighting. Sound III graphic was the most weighted for the angular and closed type, but the difference in weight from the angular and closed type was smaller. Sound four is angular and closed, but the scattering is more pronounced and occurs more frequently than the above three sounds, which is related to the clear visual characteristics produced by the percussion instruments described above. The contour impressions of the acoustic group are shown in Table 6.

Table 6: The original sound group is impressed

	No angular and closed	It has edges and corners and is closed	Open	Scattered point type
Voice 1	22	11	11	1
Voice 2	13	29	5	4
Voice 3	22	16	7	7
Voice 4	9	33	4	13

The color impressions of the acoustic group are shown in Table 7. There are six types under the dimension of color, blue color, red-pink color, yellow-orange color, green color, purple color, and no color. In order to calculate the frequency of occurrence of different colors under each sound, the calculation will be repeated when two or more colors appear in one of the sounds, so the total number of calculations will be different. The data shows that sound 1 has the highest percentage of blue color, sound 2 has the highest percentage of yellow-orange color, sound 3 has the highest percentage of green color, and sound 4 has the highest percentage of yellow-orange color.

Table 7: The original sound group color impression

	Blue system	Red system	Yellow-orange system	Green system	Purple system	Achromatic	Total
Voice 1	21	8	4	9	2	5	21
Voice 2	15	7	14	3	1	14	15
Voice 3	10	7	13	13	2	3	10
Voice 4	9	11	12	3	3	7	9

7.3 Experimental Reverse Validation and Summary

In order to verify that the emotional tendency generated by the change of visual features can match the emotional tendency generated by the sound, the contour and color impressions of the distortion and reverb groups are integrated, and the plain patterns are set up as the carriers of the texture and the dynamic effects for the additional verification of the experiment.

After data collection, the results were analyzed, and the same emotional model was used to record the findings, and the modeled images presented the emotional tendency of each dynamic pattern. It can be seen that the pattern group whose change trend is rough and chaotic is a manifestation of low pleasure and high activation (-V+A), while the pattern group whose change trend is fuzzy and planar diffusion is a manifestation of high pleasure and low activation (+V-A). The two groups of dynamic pattern emotional tendencies are shown in Figure 13.

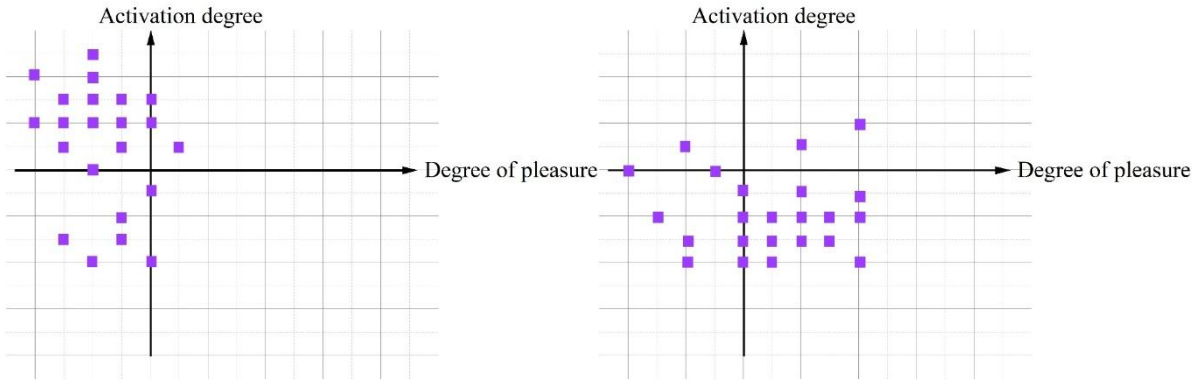


Figure 13: Two groups of dynamic pattern emotional tendencies

Then overlapping the visual emotional image with the auditory emotional image, it can be seen that the emotional model map of the distortion group has a high tendency to overlap with the emotional model map of the rough and chaotic pattern group, and the emotional model map of the reverberation group has a high tendency to overlap with the emotional model of the fuzzy and planar diffusion pattern group. The differences in the sound timbre of the aforementioned experiments bring about different emotional results, and the visual features assessed by the respective sounds, which are used to design the plain dynamic pattern, are able to cope with the emotional changes of the original sounds, and it can be assumed that the dynamic and textural texture changes of this pattern are able to stimulate the same emotional tendency of the corresponding sound samples. The sound and dynamic pattern emotional tendency overlap model is shown in Figure 14.

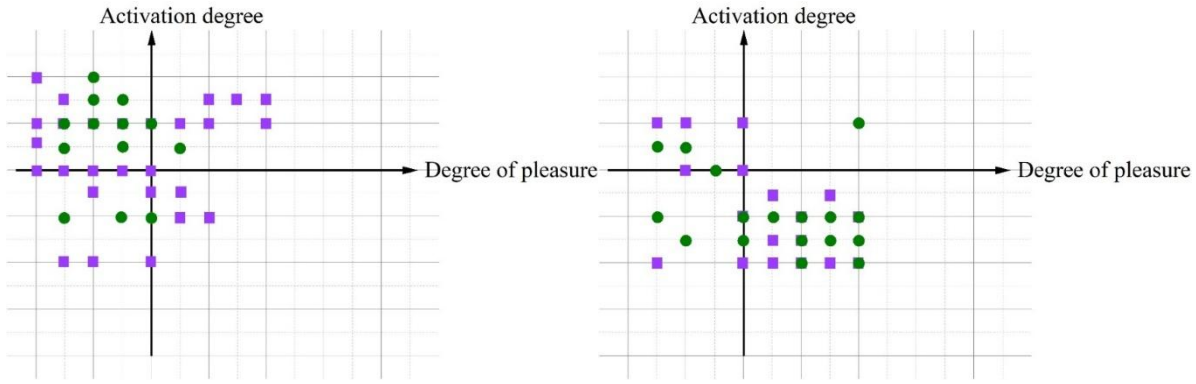


Figure 14: The model of overlapping emotional tendencies

7.4 Modeling Emotional Audiovisual Design

7.4.1 Underlying Visual Feature Construction

From the above experiments can be derived from different emotions under the auditory association tendency, from the design of the idea to refine the application of the characteristics of the method, which will help the actual use of the model, planning a good underlying foundation conducive to the results of the sound visualization of the sound is clear in a certain emotional category.

1. Color: In visual design, there is often the concept of color tone to coordinate the overall tone of the design, using color to pave the way for the basic style of emotion. Emotions can be expressed in different color intensities and can be superimposed on each other to form different moods.

2. Texture: Texture gives identity to similar parts of the color, making its role more clear, used to divide the area and enrich the visual level. The conclusion of Experiment 2 is that happy emotions are known to have soft and smooth textures, anxious emotions are known to have rough textures with a strong sense of weight, sadness is known to have both smooth and rough textures, and relaxation is known to have smooth and soft textures with a light sense of weight.

3. Kinetic effects: Experiment 2 concluded that the kinetic effects of happy emotions are mostly jumping and irregular, the kinetic effects of angry emotions are mostly impactful, the kinetic effects of sad emotions are mostly slow with a sense of resistance, and the kinetic effects of relaxing emotions are mostly more moderate and stable. In the setting of the visual characteristics of the dynamic effect, the same three elements of the movement are divided into design combinations.

4. Shape: Shape is not used to experiment and prove the output, but to set the idea through the method of cultural characteristics and sound association. The basic shapes are applied based on the characteristics of the sound emitting body and the sound associations brought by cultural influences and life experiences.

7.4.2 High-Level Abstraction Attributes and Model Generalization Summary

The overall model framework for the image semantics of the three layers, the first two layers have been narrated logic, through the bottom and middle layers of the advancement of padding, and ultimately constitute the high-level abstract attributes. From the auditory-associative visual features based on timbral characteristics, symbolic meaning and emotional preferences and other factors, with the assistance of experimental findings, set the basic auditory-associative elements. After integrating the design, the auditory-associative elements are then oriented to logical features, combined to construct images, forming object semantics, and under the comprehensive expression of different images combined with music in the context of semiotics, conveying music visualization works with more atmospheric rendering power. Through the interpretation of the picture composition object, cultural life experience influence, design techniques under the comprehensive expression, people on the sound visualization will produce more high-level abstract emotional semantics, the audiovisual design works to form their own interpretation, so that the model from the emotional response to the visual start, and then return to a richer fit of the emotional experience.

8 Conclusion

This study investigates the entropy measure based EEG signal decoding method and proposes its prediction method in auditory association. The following conclusions are drawn:

(1) Purple and blue visualized representations of music stimulate the brain more regularly compared to red.

(2) In the emotional tendency analysis, the distortion group was more agitated (-V+A), and the reverberation group was more tranquil and relaxed (+V-A), with negative correlation between the two groups, which indicates that the timbre of different effect treatments has a great influence on the emotional perception.

(3) The emotional audio-visual design model can be constructed by two levels: the bottom level visual features and the top level abstract attributes.

References

[1] Alexander, S. C., Garner, D. K., Somoroff, M., Gramling, D. J., Norton, S. A., &

Gramling, R. (2015). Using music [al] knowledge to represent expressions of emotions. Patient Education and Counseling, 98(11), 1339-1345.

[2] Davies, S. (2017). The expression of emotion in music. In Aesthetics (pp. 213-218). Routledge.

[3] Han, D., Kong, Y., Han, J., & Wang, G. (2022). A survey of music emotion recognition. Frontiers of Computer Science, 16(6), 166335.

[4] Yang, X., Dong, Y., & Li, J. (2018). Review of data features-based music emotion recognition methods. Multimedia systems, 24(4), 365-389.

[5] Aljanaki, A., Yang, Y. H., & Soleymani, M. (2017). Developing a benchmark for emotional analysis of music. PloS one, 12(3), e0173392.

[6] Soleymani, M., Aljanaki, A., Yang, Y. H., Caro, M. N., Eyben, F., Markov, K., ... & Wiering, F. (2014, November). Emotional analysis of music: A comparison of methods. In Proceedings of the 22nd ACM international conference on Multimedia (pp. 1161-1164).

[7] Cui, X., Wu, Y., Wu, J., You, Z., Xiahou, J., & Ouyang, M. (2022). A review: Music-emotion recognition and analysis based on EEG signals. Frontiers in neuroinformatics, 16, 997282.

[8] Zhou, T. H., Liang, W., Liu, H., Wang, L., Ryu, K. H., & Nam, K. W. (2022). EEG emotion recognition applied to the effect analysis of music on emotion changes in psychological healthcare. International Journal of Environmental Research and Public Health, 20(1), 378.

[9] Panda, R., Malheiro, R., & Paiva, R. P. (2020). Audio features for music emotion recognition: a survey. IEEE Transactions on Affective Computing, 14(1), 68-88.

[10] Choi, J., Song, J. H., & Kim, Y. (2018, June). An analysis of music lyrics by measuring the distance of emotion and sentiment. In 2018 19th IEEE/ACIS international conference on software engineering, artificial intelligence, networking and parallel/distributed computing (SNPD) (pp. 176-181). IEEE.

[11] Garima, Goel, N., & Rathee, N. (2023). Modified multidimensional scaling on EEG signals for emotion classification. Multimedia Tools and Applications, 82(18), 28547-28568.

[12] Chen, K., Liu, Q., Ai, Q., Zhou, Z., Xie, S. Q., & Meng, W. (2016). A MUSIC-based method for SSVEP signal processing. Australasian physical & engineering sciences in medicine, 39(1), 71-84.

[13] Han, O., & Kim, J. (2020). Analysis and Process of Music Signals to Generate Two-Dimensional Tabular Data and a New Music. Computers, Materials & Continua, 63(2).

[14] Kountchev, R., Mironov, R., Draganov, I., & Kountcheva, R. (2022, July). Multidimensional Signal Processing and Applications—New Approaches. In International Workshop on New Approaches for Multidimensional Signal Processing (pp. 3-9). Singapore: Springer Nature Singapore.

[15] Madanayake, A., Wijenayake, C., Lin, Z., & Dornback, N. (2015, May). Recent advances in multidimensional systems and signal processing: An overview. In 2015 IEEE International Symposium on Circuits and Systems (ISCAS) (pp. 2365-2368). IEEE.

[16] Shi, Y. Q., Su, W., Chen, C. M., Rajala, S. A., Bose, N. K., & Sibul, L. H. (2018). Multidimensional Signal Processing. In Circuits, Signals, and Speech and Image Processing (pp. 18-1). CRC Press.

[17] Bourennane, S., Marot, J., Fossati, C., Bouridane, A., & Spinnler, K. (2014). Multidimensional signal processing and applications. The Scientific World Journal, 2014, 1-2.

[18] Rabenstein, R., & Velten, J. (2014). Special issue on multidimensional signal processing applications. Multidimensional Systems and Signal Processing, 25(2), 245-246.

[19] Dhanasingh, A., & Hochmair, I. (2021). Signal processing & audio processors. Acta Oto-Laryngologica, 141(sup1), 106-134.

[20] Purwins, H., Li, B., Virtanen, T., Schlüter, J., Chang, S. Y., & Sainath, T. (2019). Deep learning for audio signal processing. IEEE Journal of Selected Topics in Signal Processing, 13(2), 206-219.

[21] Costantini, G., Casali, D., & Cesarini, V. (2024). New advances in audio signal processing. Applied Sciences, 14(6), 2321.

[22] Santosh, K. C., Borra, S., Joshi, A., & Dey, N. (2019). Preface: special section: advances in speech, music and audio signal processing (articles 1–13). International journal of speech technology, 22(2), 293-294.

[23] Gómez-Cañón, J. S., Cano, E., Eerola, T., Herrera, P., Hu, X., Yang, Y. H., & Gómez, E. (2021). Music emotion recognition: Toward new, robust standards in personalized and context-sensitive applications. IEEE Signal Processing Magazine, 38(6), 106-114.

[24] Aljalal Majid,Aldosari Saeed A.,Molinas Marta,AlSharabi Khalil & Alturki Fahd A.. (2022). Detection of Parkinson's disease from EEG signals using discrete wavelet transform, different entropy measures, and machine learning techniques. Scientific Reports,12(1).

[25] Yao Longxin,Wang Mingjiang,Lu Yun,Li Heng & Zhang Xue. (2021) .EEG-Based Emotion Recognition by Exploiting Fused Network Entropy Measures of Complex Networks across Subjects. Entropy,23(8),984-984.

[26] Ajay Kumar Maddirala & Rafi Ahamed Shaik. (2016). Motion artifact removal from single channel electroencephalogram signals using singular spectrum analysis. Biomedical Signal Processing and Control,30,79-85.