Mixture Kernel Density Estimation and Remedied Correlation Matrix on the EEG-Based Copula Model for the Assessment of Visual Discomfort

Yawen Zheng[1] · Xiaojie Zhao[1] · Li Yao[1]

Abstract

Since electroencephalogram (EEG) signals can directly provide information on changes in brain activity due to behaviour changes, how to assess visual discomfort through EEG signals attracts researchers' attention. However, previous assessments based on time-domain EEG features lack sufficient consideration of the dependence among EEG signals, which may affect the discrimination to visual discomfort. Although the copula model can explore the dependence among variables, the EEG-based copula models still have the following deficiencies: (1) the methods ignoring the fine-grained information hidden in EEG signals could make the estimated marginal density function improper, and (2) the approaches neglecting the pseudo-correlation among data may inappropriately estimate the correlation matrix parameter of the copula density function. The mixture kernel density estimation (MKDE) and remedied correlation matrix (RCM) on the EEG-based copula model are proposed to mitigate the mentioned shortcomings. The simulation experiments show that MKDE can not only better estimate the marginal density function but also explore fine-grained information. The RCM can be closer to the real correlation matrix parameter. With the favourable quality of the proposed EEG-based model, it is used to extract time-domain EEG features to assess visual discomfort further. To our best knowledge, the extracted features present better discrimination to visual discomfort compared with the features extracted by the state-of-the-art method.

Keywords Mixture kernel density function · Remedied correlation matrix · EEG-based copula model · Time-domain EEG features · Visual discomfort

Introduction

Stereoscopic displays are prevalent in our daily lives, such as entertainment [1, 2], education [3] and medical treatment [4]. However, the further development of stereoscopic displays is limited by visual discomfort induced by stereoscopic devices [5]. The electroencephalogram (EEG) signals can not only directly provide real-time brain activities, but also

have high-quality temporal resolution. Therefore, researchers try to assess visual discomfort through EEG signals and find the features closely related to visual discomfort [6–8].

The assessment quality highly depends on the extracted features from EEG signals. According to the research of Monteiro et al. [9], the methods used to extract the features in the field of assessing discomfort include power spectral density (PSD) [10–12], statistical analysis, entropy measures, etc. For methods to extract the time-domain EEG features, statistical analysis is generally used with event-related potential (ERP) signals [13, 14], calculated from EEG signals. However, the strictly millisecond duration of the target stimulus in the ERP experiment may not induce visual discomfort. Thereby, researchers turn to other methods. Recently, Peng et al. [15] have proposed a multiscale entropy, a state-of-the-art method, to extract features highly related to visual discomfort caused by stimuli with different flickering frequencies. With support vector machine (SVM) as the classifier, the extracted features are more sensitive

✉ Li Yao
yaoli@bnu.edu.cn

Yawen Zheng
linyazyw@mail.bnu.edu.cn

Xiaojie Zhao
zhaox86@163.com

[1] School of Artificial Intelligence, Engineering Research Center of Intelligent Technology and Educational Application, Beijing Normal University, Beijing, China

to visual discomfort than the features extracted by sample entropy that is one of the conventional entropy measures.

Admittedly, the current methods have made progress in extracting the time-domain EEG features related to visual discomfort, whereas these methods ignoring the inter-signal dependence may affect the discrimination of existing features for visual discomfort. Notably, the copula model is considered powerful to explore the dependence among variables, including the structure of dependence. It connects the marginal density functions with the copula density function to restore the joint density function that implicitly presents the dependence among variables [16]. Because the copula model can revive the dependence without assumptions of distributions or constraints on variables, it has been widely used in analysing the times series that being with strong dependence or complex dependence structures [17–19]. It is known that EEG signals from different channels could present complex dependence. Inspired by it, some researchers try to introduce the copula model in the analysis of EEG signals.

The key to establishing the EEG-based copula model (copula model for analysing EEG signals) is to estimate the marginal density functions and the parameters of the copula density function. Specifically, Iyengar et al. [20] used the EEG-based copula model to quantify the synchrony of EEG signals for assisting in the early diagnosis of Alzheimer's disease, in terms of the high correlation among synchronized EEG signals. The multi-information and Kullback–Leibler (KL) divergence were introduced to estimate the correlation matrix and the degree of freedom parameters of the copula density function by finding the narrowest gap between the real and the estimated copula density function. The accuracy of the proposed EEG-based copula model to early diagnose Alzheimer's disease presented 2% higher than conventional models. Qian et al. [21] proposed an EEG-based copula Bayesian classifier to detect drowsiness during the short daytime nap, according to drowsiness signals presenting the different dependence from the alertness signals. The kernel density estimation (KDE) and model selection criteria (Akaike information criterion (AIC) and Bayesian information criterion (BIC)) were used separately to estimate marginal density functions and the correlation matrix parameter of the copula density function. The accuracy of the proposed classifier for drowsiness detection was 94.3%, which performed best among the list models. Fontaine et al. [22] adopted the EEG-based copula model to probe the changepoints of EEG signals for analysing the impact of stroke on the local field potential of rats, based on that the dependence of pre-changepoints would be different from the post-changepoints. The BIC was introduced to estimate marginal density functions. The model selection criteria (AIC, BIC, and correlation information criterion (CIC)) and Kullback–Leibler information criteria (KLIC) were applied to estimate the correlation matrix and the degrees of freedom parameters of the copula density function. The Kolmogorov–Smirnov test showed that the parameters of EEG-based copula on pre-changepoints were significantly different from that on post-changepoints. In other words, the proposed model could probe changepoints in the local field potentials of rats.

According to the previous studies, the EEG-based copula model is only in its infancy. Therefore, there is still room to improve the EEG-based copula model. For example, the KDE used to estimate marginal density functions with overall data may ignore potentially fine-grained information, which may cause a gap between the estimated and the real marginal density functions. Without considering the pseudo-correlation among EEG signals caused by noise, distortion, etc., the parameters of the copula density function estimated by model criteria may deviate from the real parameters of the copula density function. Inspired by it, the mixture kernel density estimation (MKDE) and remedied correlation matrix (RCM) on the EEG-based copula model are proposed. The MKDE is used to estimate the marginal density function of the overall data by a mixture of the estimated marginal density function of each cluster clustered by K-means. The marginal density function of each cluster is estimated by KDE incorporated with the multiscale strategy. The MKDE cannot only better estimate the marginal density function but also explore the fine-grained information. The RCM is estimated by correlation analysis of the remedied EEG data equalling to the sparse matrix obtained by locality-sensitive dictionary learning. In terms of dictionary learning, the remedied EEG data can use fewer elements of the dictionary matrix to restore the original EEG data. In other words, the remedied EEG data learning most of the original data is supposed to contain less pseudo-correlation. The RCM may be closer to the real correlation matrix parameter of the copula density function.

The main contributions are outlined as:

– The MKDE employing the KDE combined with the K-means and the multiscale strategy can explore the fine-grained information hidden in the original data and better estimate the marginal density function.
– The RCM estimated by correlation analysis based on the remedied data obtained by the dictionary learning may contain less pseudo-dependence, which could be closer to the real correlation matrix parameter of the copula density function.
– The proposed EEG-based copula model is used to assess visual discomfort induced by the stereoscopic display. Compared with the time-domain EEG features extracted by the state-of-the-art method, the features extracted by our proposed model present better discrimination to visual discomfort.

This paper is organized as follows. Section 2 briefly covers the review of the theory of the copula model. Section 3 explains the improved EEG-based copula model. Section 4 shows the evaluation of the improved EEG-based copula model, and experimental results and discussions are reported in Section 5 and Section 6, respectively. Section 7 gives the conclusions.

Primitive Copula Model

According to the theory of Sklar [23], the joint probability distribution of multivariate can be restored by a copula function with its marginal probability distributions as variables. The copula function is as follows:

$$F(x_1, x_2, ..., x_N) = C(F_1(x_1), F_2(x_2), ..., F_N(x_N)) \quad (1)$$

where $F(\cdot)$ is the joint probability distribution, $x_i (i \in [1, N])$ represents the variable, $C(\cdot)$ is the copula function, and $F_i(\cdot)$ $(i \in [1, N])$ is the marginal probability distribution of each variable.

If the joint probability distribution $F(\cdot)$ is continuous, and the marginal probability distribution $F_i(\cdot)$ is strictly increasing and continuous, equation (1) can be transformed to:

$$f(x_1, x_2, ..., x_N) = c(F_1(x_1), F_2(x_2), ..., F_N(x_N)) \prod_{i=1}^{N} f_i(x_i) \quad (2)$$

where $f(\cdot)$ is the joint density function, $c(\cdot)$ is the copula density function, and $f_i(\cdot)$ $(i \in [1, N])$ represents the marginal density function of each variable.

Equation (2) indicates that the joint density function can be restored by the product of the marginal density functions and the copula density function. The marginal density function of each variable does not have to follow the same distribution. The choice of copula density function depends neither on the marginal density functions nor on the marginal probability function used as a variable. Better than traditional correlation analysis, such as Pearson correlation and Spearman correlation, the copula density function can estimate the correlation relationship without any restrictions on the distribution of variables.

Improved EEG-Based Copula Model

In this section, the proposed MKDE and RCM on the EEG-based copula model are illustrated in detail. As Fig. 1 shows, the MKDE and RCM aim to estimate the marginal density function and be closer to the real correlation matrix of the copula density function. There is no interference between the MKDE and RCM.
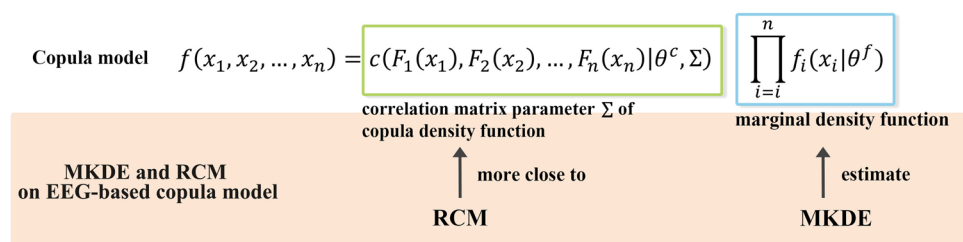
Mixture Kernel Density Estimation

The marginal density function of an arbitrary variate can be estimated by KDE with smooth kernel functions. Based on Scott's research [24], the quality of KDE greatly depends on the smoothing parameters of the kernel functions. Assuming there are N d−dimensional samples $X = [x_1, x_2, ..., x_N]$, the marginal density function of the variate X can be smoothly estimated by:

$$\widehat{f}_h(x) = \frac{1}{N} \sum_{i=1}^{N} K_H(x - x_i) \quad (3)$$

where $K_H(\cdot)$ is the non-negative kernel function. The H is $d \times d$ symmetric and positive definite matrix which determines the smoothness of the estimated marginal density function [25]. Once the d equals to 1, the matrix H will degenerate to a scalar h.

As the number and the dimension of samples increase, the complexity and the cost of KDE can be relatively high. Therefore, KDE combined with binning strategy [26] or with K-means [27], reduced set density estimator [28], finite mixture model [29], etc., are proposed to reduce the scale of data to mitigate the mentioned problem. However, these methods or models ignore the fine-grained information hidden in the overall samples, affecting the estimated marginal density function. Based on the assumption that the overall data may be composed of several subcomponents [25, 27], the MKDE employing the KDE with the K-means and multiscale strategy is proposed to explore fine-grained information for better estimation. Concretely, the whole data are firstly clustered into several clusters by K-means, which is a classic unsupervised method in the field of clustering. After that, to dig out the fine-grained information, KDE with different smoothing parameters, hereinafter multiscale strategy,

Fig. 1 The illustration of the proposed EEG-based copula model

Copula model $\quad f(x_1, x_2, ..., x_n) = \boxed{c(F_1(x_1), F_2(x_2), ..., F_n(x_n) | \theta^c, \Sigma)} \quad \boxed{\prod_{i=i}^{n} f_i(x_i | \theta^f)}$

correlation matrix parameter Σ of copula density function   marginal density function

MKDE and RCM on EEG-based copula model   more close to   estimate

RCM   **MKDE**

is used to estimate the marginal density function of each cluster. In terms of the theory of the KDE, the larger and smaller smoothing parameters separately focus on global data and regional data. Therefore, with the multiscale strategy, smaller smoothing parameters can explore fine-grained information. Meanwhile, the other larger smoothing parameters can compensate for the non-smoothness of estimated marginal density function caused by the smaller smoothing parameters. The concrete equations are as follows:

$$f^*(X) = \sum_{k=1}^{C} \lambda_k \widehat{f}_k(x) \tag{4}$$

where $f^*(\cdot)$ represents the marginal density function of the N 1−dimensional samples data $X = [x_1, x_2, ..., x_N]$. The k is the kth cluster in C clusters. The coefficient λ_k represents the contribution of the marginal density function of the kth cluster to the marginal density function of the overall data, which is the ratio of the number of data in the kth cluster to the entire data. The $\widehat{f}_k(\cdot)$ is the estimated marginal density function of the kth cluster.

$$\widehat{f}_k(x) = \sum_{j=1}^{m} \alpha_j \widehat{f}_j(x) \tag{5}$$

where j means the jth smoothing parameter among m smoothing parameters. The coefficient α_j represents the contribution of the jth smoothing parameter to the estimated marginal density function of the kth cluster, which is the reciprocal of m smoothing parameters. The $\widehat{f}_j(\cdot)$ is the marginal density function estimated by KDE with the jth smoothing parameter.

$$\widehat{f}_j(x) = \frac{1}{N_k} \sum_{i=1}^{N_k} K_{h_j}\left(\frac{x - x_i}{h_j}\right) \tag{6}$$

where N_k represents the number of data in the kth cluster. And x_i is the ith data in the kth cluster. The h_j and $K_{h_j}(\cdot)$

are the value of and the kernel function of the jth smoothing parameter.

The framework and procedure of MKDE are, respectively, shown in Fig. 2 and Alg. 1. As Fig. 2 shows, the overall EEG data are grouped into several clusters by K-means. The candidates for the number of clusters should be set in advance. The GridSearch algorithm is introduced to find the best result, a general exhaustive method for finding the optimal parameters. Although the smaller smoothing parameter can explore fine-grained information, too small parameters may be easily disturbed by noise. Thereby, the GridSearch algorithm is also used to find the optimal number of smoothing parameters. With different smoothing parameters, the estimated marginal density function of a single cluster may present some differences, as the dashed blue rectangle in Estimation with multiscale part shows. The final marginal density function of a single cluster is obtained by the weighted summation of the estimated marginal density function by different smoothing parameters, as the dashed blue rectangle in Estimation of each cluster part shows. The estimated marginal density function of the overall data is weighted mixed by the estimation results of clusters, as the final estimation with cluster part depicts.

Remedied Correlation Matrix

There are several candidates for the copula density function, such as Gaussian copula, Student's t copula, and Clayton copula [30–32]. The Gaussian copula is selected in this paper for its wide applications on time series. Each variate in the Gaussian copula density function is assumed to follow the Gaussian distribution for simplifying the Gaussian copula density function. The Gaussian copula density function is expressed as:

$$c(X) = c_{\mathcal{N}(\mu,\Sigma)}(F_{\mathcal{N}(\mu_1,\Sigma_1)}(x_1), F_{\mathcal{N}(\mu_2,\Sigma_2)}(x_2), ..., \\ F_{\mathcal{N}(\mu_N,\Sigma_N)}(x_N)) \tag{7}$$

Fig. 2 The framework of the MKDE on the EEG-based copula model

---
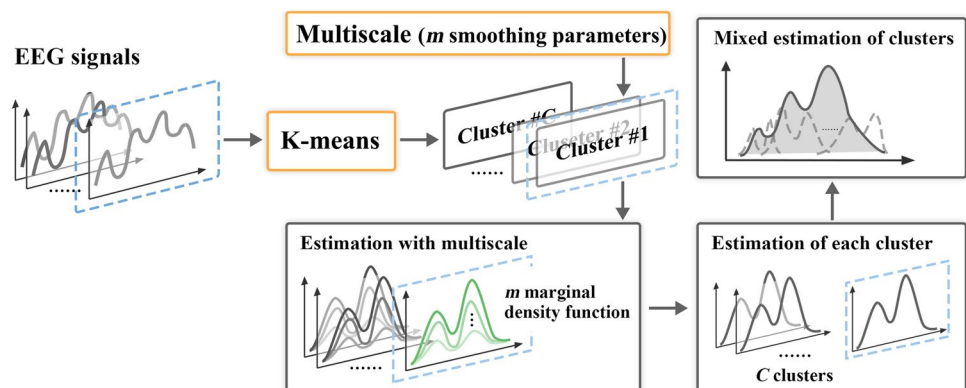
Algorithm l: MKDE

Input:
input data (X); C_s (candidates for the number of clusters); m (number of smoothing parameters);
h_init (range of smoothing parameters);
Output:
h (set of smoothing parameters); $f^*(X)$ (estimated marginal density function);
1: $C \leftarrow \text{GridSearch}(C_s, K\_means, X)$;
2: x (clustered data) $= K\_means(X, C)$;
3: $summation\_f\_k = 0$;
4: **for** $k = 1$ to C **do**
5: $\quad \lambda_k = N_k$ (number of x) / N (number of X);
6: $\quad h^k_{best}(m) \leftarrow \text{GridSearch}(x, \text{KDE}, h\_init)$
7: $\quad h_k \leftarrow h^k_{best}(m)$;
8: $\quad \widehat{f_k}(x) = 0$;
9: $\quad$ **for** $j = 1$ to m **do**
10: $\quad\quad f_j(x) \leftarrow \text{KDE}(h^k_{best}(j))$;
11: $\quad\quad \alpha_j = 1 / m$;
12: $\quad\quad \widehat{f_k}(x) = \widehat{f_k}(x) + \alpha_j * f_j(x)$;
13: $\quad$ **end**
14: $\quad summation\_f\_k = summation\_f\_k + \lambda_k * \widehat{f_k}(x)$;
15: **end**
16: $f^*(X) = summation\_f\_k$;
17: $h \leftarrow [h_1, h_2, ..., h_C]$.

---

where $c(\cdot)$ is the Gaussian copula density function. The parameters μ and Σ are separately the mean and the correlation matrix of $c(\cdot)$. The parameters μ_i and $\Sigma_i (i = 1, ..., N)$ are the mean and the correlation matrix of the Gaussian distribution of variate x_i.

Because this paper focuses on better extracting the time-domain EEG features by the EEG-based copula model, estimating the correlation matrix Σ is crucial. The locality-sensitive dictionary learning is introduced to mitigate the pseudo-correlation among EEG data caused by noise, distortion, etc., for better estimating the correlation matrix. Specifically, based on the theory of the dictionary learning, the original data can be decomposed into a dictionary matrix that learns the essence of the original data, and a sparse matrix (hereinafter remedied data) that learns the most important information of the original data [33, 34]. Because the remedied data can keep the most relevant information to original data, it is considered to contain less pseudo-correlation. In the conventional dictionary learning, only considering the constraints on coefficients, locality-sensitive dictionary learning has the advantage of exploring the local information of data for data construction [35–39]. Therefore, locality-sensitive dictionary learning is used to obtain remedied data. The columns in the remedied data are considered as variables to estimate RCM. As equation (7) indicates, the marginal probability distribution of each variable x_i is required to be estimated. Assuming that each variable x_i follows a Gaussian distribution, the parameters

μ_i and Σ_i can be obtained from the sample mean and sample variance. The data sampled from the estimated Gaussian distribution are concatenated in the order of the columns of the remedied data. The RCM is finally estimated by correlation analysis of the concatenated data. The process of estimating the RCM is shown in Fig. 3.

Evaluation of the Improved EEG-based Copula Model

The improved EEG-based copula model is evaluated by simulated EEG data and real EEG data, respectively. Due to a lack of public EEG datasets on visual discomfort, two experiments (the behaviour experiment and the EEG experiment) are elaborately designed to collect real EEG data. The framework of assessing visual discomfort by the proposed model is illustrated at last.

Simulation Experiment

Five kinds of simulated EEG data are generated to evaluate MKDE, namely clean data, data with signal to noise (SNR) of 5dB, data with SNR of 10dB, data with SNR of 15dB, data with SNR of 20dB. Simulated EEG data induced by auditory and visual stimuli are generated to evaluate RCM. All simulated EEG data are generated by the MNE-Python package.

Fig. 3 The framework of the RCM on the EEG-based copula model
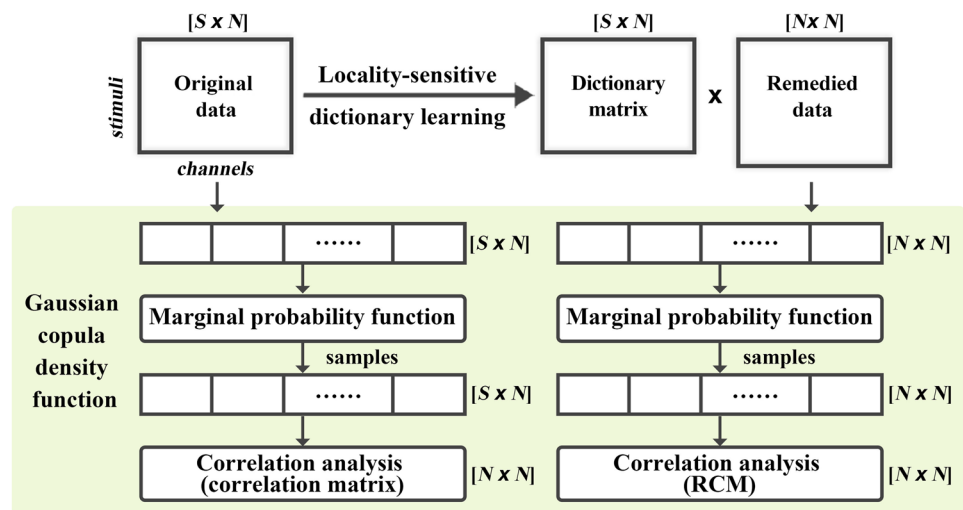

Behaviour Experiment

The behaviour experiment aims to subjectively assess visual discomfort induced by the designed stereoscopic stimuli to assist in selecting reasonable stimuli that can be used in the EEG experiment. In terms of previous studies [7, 40], the symptom questionnaire is chosen to be the subjective questionnaire. The question in the subjective questionnaire is whether the stimulus that just appeared makes you visual discomfort. Participants need to select one from the five choices (no fatigue, mild fatigue, moderate fatigue, serious fatigue, and extreme fatigue) to rate the degree of visual discomfort according to their feelings. These five choices are, in turn, quantified as five levels (1, 2, 3, 4, 5). The rounded average level from all participants is regarded as a label of the stimulus. Referring to the studies of other researchers and our previous research [5, 40–42], disparity images with disparities ranging from ±0.1° to ±1° are used as stimuli. The designed stimuli are presented by the LG D2343P

display with 1920×1080 resolution. Seventeen participants aged from 23 to 27 with normal stereo vision were recruited from Beijing Normal University.

EEG Experiment

The EEG experiment is designed to record the EEG signals of participants when facing stimuli for objectively assessing visual discomfort. With the results of the behaviour experiment, ten disparity images are selected that can effectively induce visual discomfort or maintain visual comfort. Forty-three participants were recruited from Beijing Normal University, including 22 males and 21 females. Before the experiment, each participant was suggested to stay healthy, avoid staying up late and excessive stress. The EEG signals were recorded by the 128-channel EGI system. The sampling frequency is 500 Hz. Each participant has a chance to practice several times before the formal experiment starts. In Fig. 4, the 'Cross' image is presented to fix the participant's vision. Once the

Fig. 4 The procedure of the EEG experiment
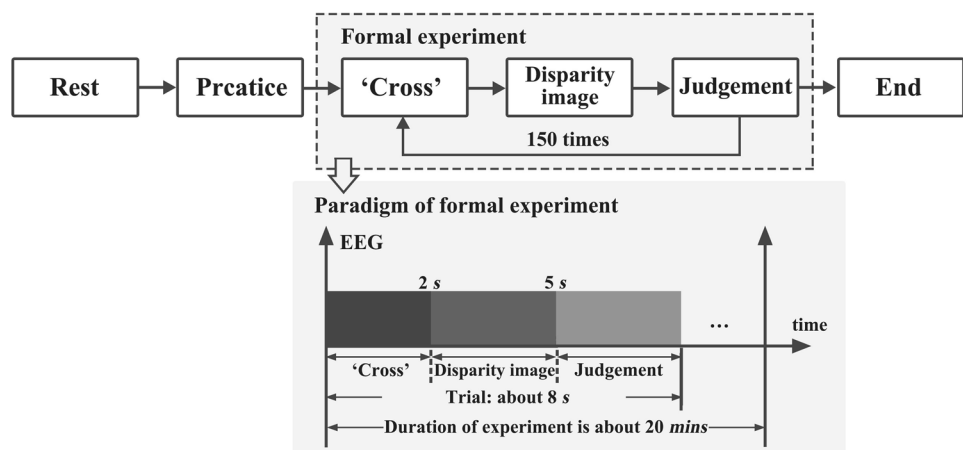
stimulus disappears, the judgment image will be presented to encourage each participant to determine whether the stimulus just appeared in front of or behind the screen. The judgment part aims at keeping the participant's attention on the current experiment. The repetition times are set to 150 to balance the trade-off between recording the EEG data as much as possible and preventing the long experimental time from inducing extra visual discomfort.

Framework of Assessing Visual Discomfort

Because raw EEG data are easily disturbed by irrelevant components, such as eye movement and head movement, independent component analysis (ICA) is introduced to eliminate the first two or three components.

Inspired by researches [43, 44], the autoregressive model (AR model) is introduced to extract rough EEG features, one of the classic methods for roughly extracting time-domain EEG features. The coefficients of the AR model (the extracted rough EEG features) are determined by AIC. For each participant, since the dimension of rough EEG features is high ([number of channels × number of AR coefficients]), the dimension reduction is required for reducing the computation cost. The Random Forest is adopted as it not only keeps valuable features but also shows the importance of these features [45]. Concretely, the importance of channels can be ranked by Random Forest following the Gini coefficient when channels of rough EEG features are considered as the nodes of Random Forest. Based on the results of all participants, the top 40 most important and frequent channels are kept. The reduced dimension of rough EEG features is [number of selected channels × number of AR coefficients].

As Fig. 5 presents, the extracted EEG features by our proposed model are the linear concatenation of the smoothing parameters from MKDE and the RCM according to the order of selected channels. The smoothing parameters can be obtained by MKDE used to estimate the marginal density function of the overall data from an individual channel. The RCM is obtained by the remedied data from a specific stimulus. Based on the extracted EEG features, SVM with radius basis function (RBF) kernel is used to assess visual discomfort. The parameters C and gamma of SVM are determined by the GridSearch algorithm.

Experimental Results

Simulation Experiments on MKDE

As Fig. 6(a) shows, with the SNR increases, simulated EEG data are gradually close to clean data. In Fig. 6(b), the grey, green and blue lines individually represent the real marginal density function, the estimated marginal density function by KDE and the estimated marginal density function by MKDE. The estimation results of the clean data or the data with high SNRs by MKDE are much closer to the real marginal density functions. However, the estimation results of the heavily contaminated data (as the second and third subplots in Fig. 6(a) presents) by MKDE deviate from the real marginal density functions.

The mean-squared error (MSE) is used to evaluate the differences between the real and the estimated marginal density functions for objectively testifying MKDE. In Fig. 7, the grey and blue lines separately represent the estimation results by KDE and MKDE. As SNR increases, the MSE of the estimation by both methods is gradually stable. However, the MSE of the estimation by MKDE is smaller. For heavily contaminated data, the gap between the performance of MKDE and the performance of KDE is narrowed. But if the number of clusters is set to one, the performance of MKDE is the same as KDE.

The details of MKDE exploring the fine-grained information hidden in the overall data are shown in Fig. 8. Better than KDE with only one smoothing parameter, MKDE not

Fig. 5 The framework of assessing visual discomfort by the proposed model

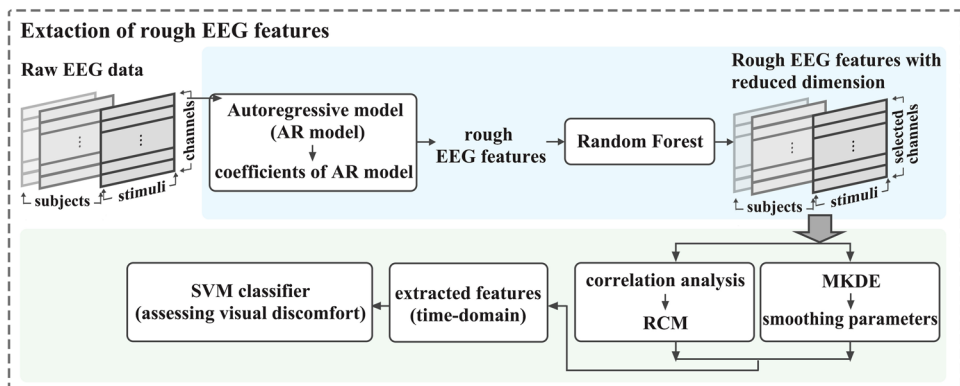The assessment of visual discomfort by the proposed model

Fig. 6 The results of real and estimated marginal density functions of simulated EEG data: (**a**) the time-domain diagrams of clean data, data with SNR of 5dB, data with SNR of 10dB, data with SNR of 15dB, data with SNR of 20dB from left to right; (**b**) the real marginal density function, estimated marginal density function by KDE and estimated marginal density function by MKDE corresponding to (**a**)
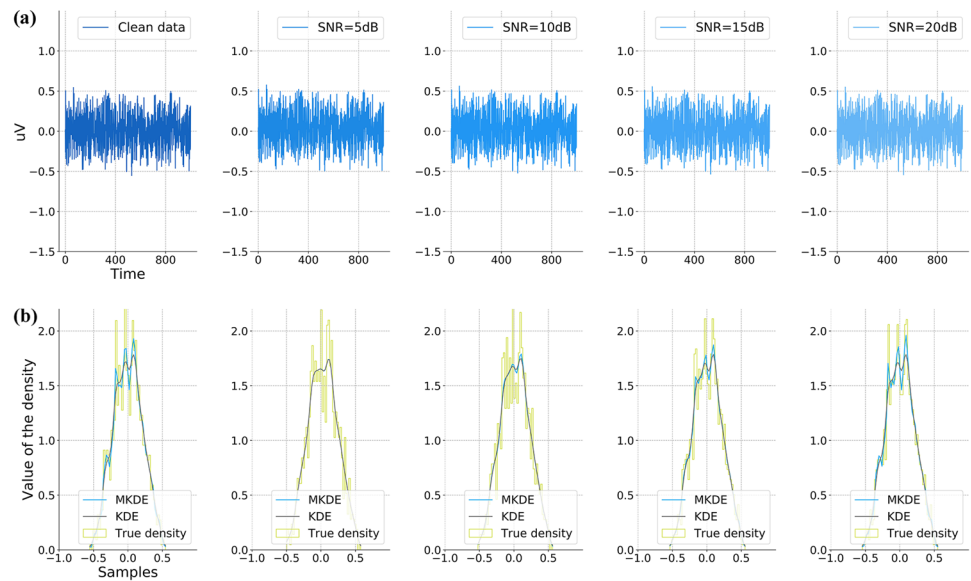
only uses the best smoothing parameter to estimate the marginal density function but also employs several smoothing parameters to ensure that hidden information at different fine-grained can be explored. Fig. 8(c) shows that the marginal density function of the data from one cluster estimated by selected smoothing parameters present apparent differences, which further indicates that these parameters can mine the fine and different-grained information. The marginal density function of one cluster, showing in Fig. 8(d), is obtained by the weighted summation of the estimation with selected smoothing parameters. Repeating the process presented in Fig. 8(c) and Fig. 8(d), the marginal density function of each cluster can be estimated, as Fig. 8(e) shows. The final marginal density function of the overall data is estimated by a weighted mixture of the estimated results from each cluster.

Simulation Experiments on RCM

Figure 9 shows the heatmap results of the correlation matrix parameters of the copula density functions obtained by original data and remedied data. The presented heatmap results are randomly selected from the heatmap results of all correlation matrices. In Fig. 9(a) and 9(b), the left and the right subfigures present correlation matrices obtained by the original data and the remedied data. According to Chen et al. [46], channels in the fronto-central region and occipital region are highly sensitive to hearing and vision,

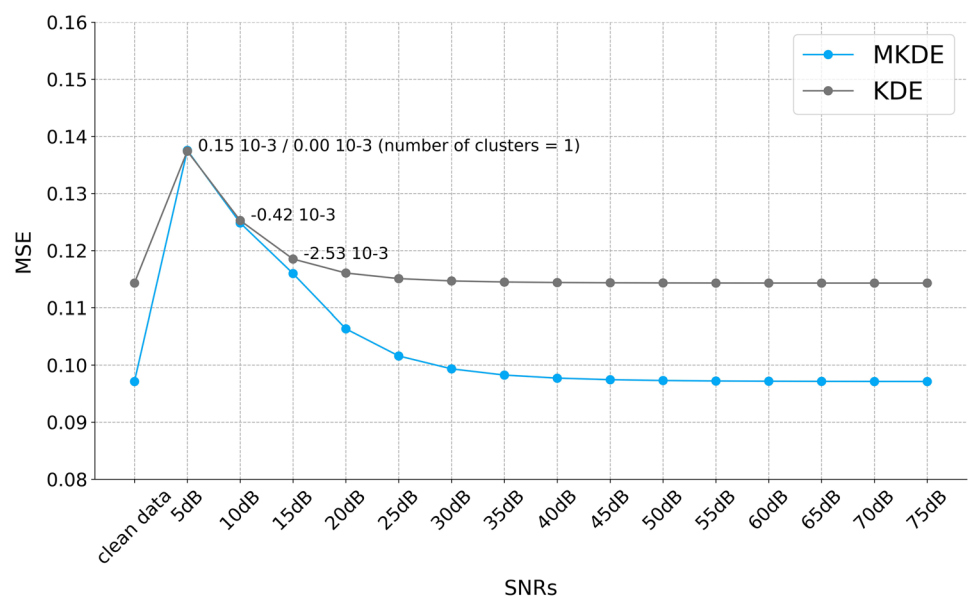**Fig. 7** The MSE of estimated marginal density functions by KDE and MKDE

Fig. 8 The details of MKDE exploring the fine-grained information: (**a**) the time-domain diagram of simulated EEG data with SNR of 50dB; (**b**) the clustering result of simulated EEG data; (**c**) the marginal density function of a single cluster, respectively, estimated with different smoothing parameters. *Sp* represents the smoothing parameter; (**d**) the marginal density function of a single cluster estimated by weighted summation of estimated results with smoothing parameters; (**e**) the estimated marginal density function of each cluster; (**f**) the estimated marginal density function of the overall data
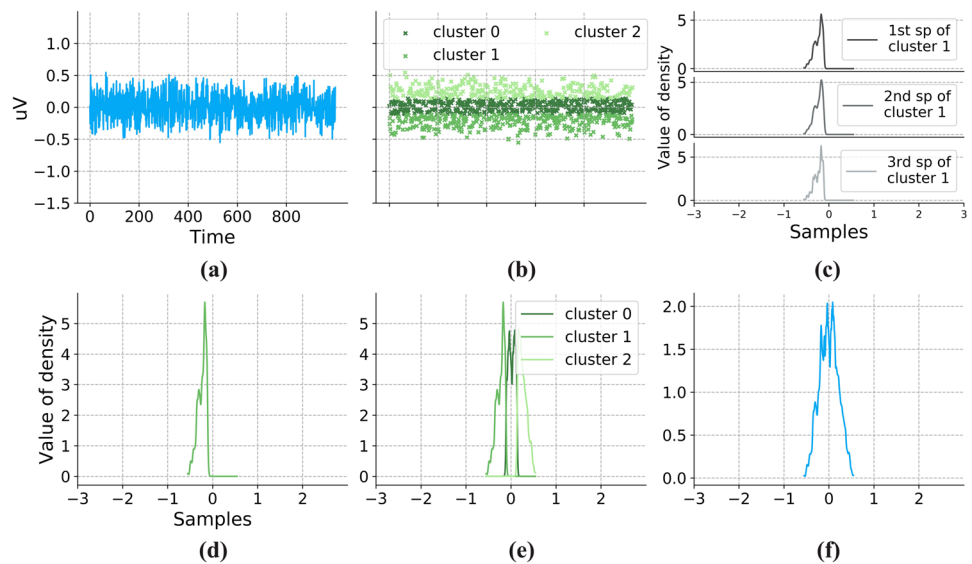
respectively. Hence, channel Fz is selected to be an example of further analysis because it is sensitive to the auditory stimulus. As the white rectangle in Fig. 9(a) shows, the correlation value between the channels FC1 and Fz, and channels C4 and Fz in RCM gets increased. Both correlation values between channels T7 and Fz, and channels P3 and Fz in RCM get decreased. The correlation relationship between channels O1 and Fz in RCM remains stable. All findings

Fig. 9 The heatmap results of correlation matrices of the copula density function: (**a**) the heatmap results of correlation matrices obtained by simulated original data and remedied data corresponding to the 11th auditory stimulus; (**b**) the heatmap results of correlation matrices obtained by simulated original data and remedied data corresponding to the first visual stimulus
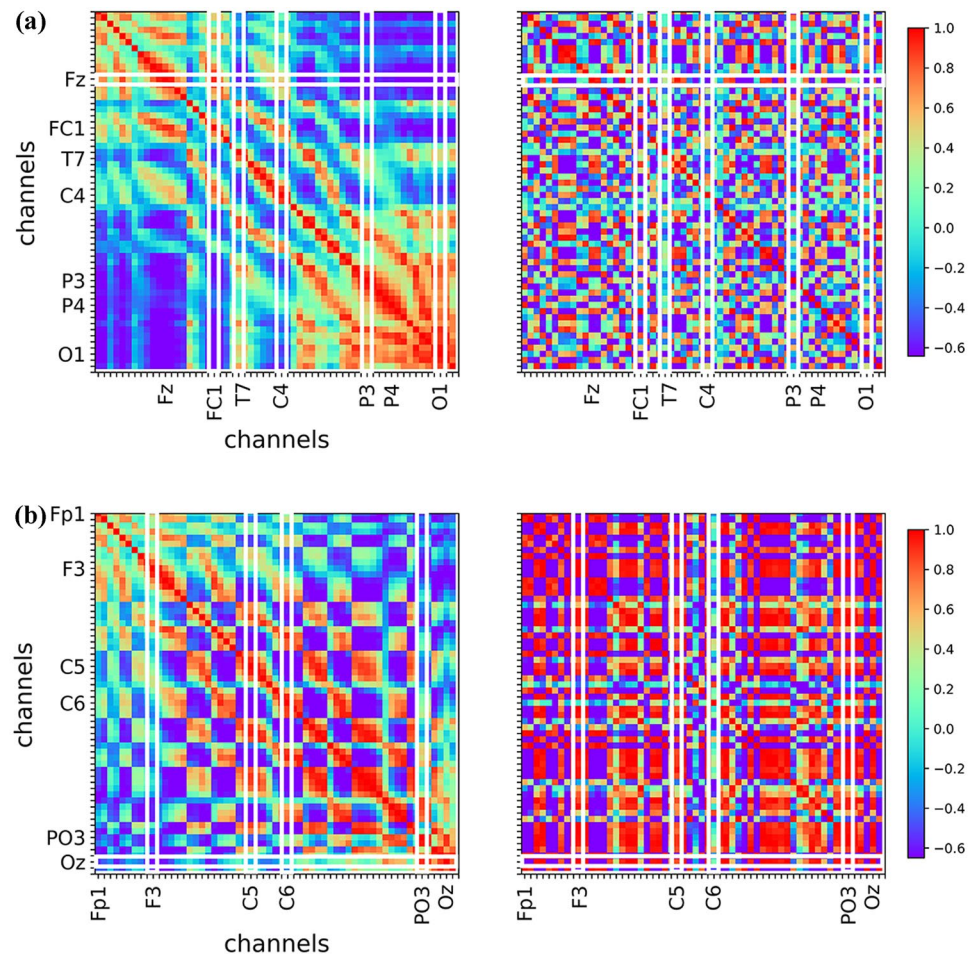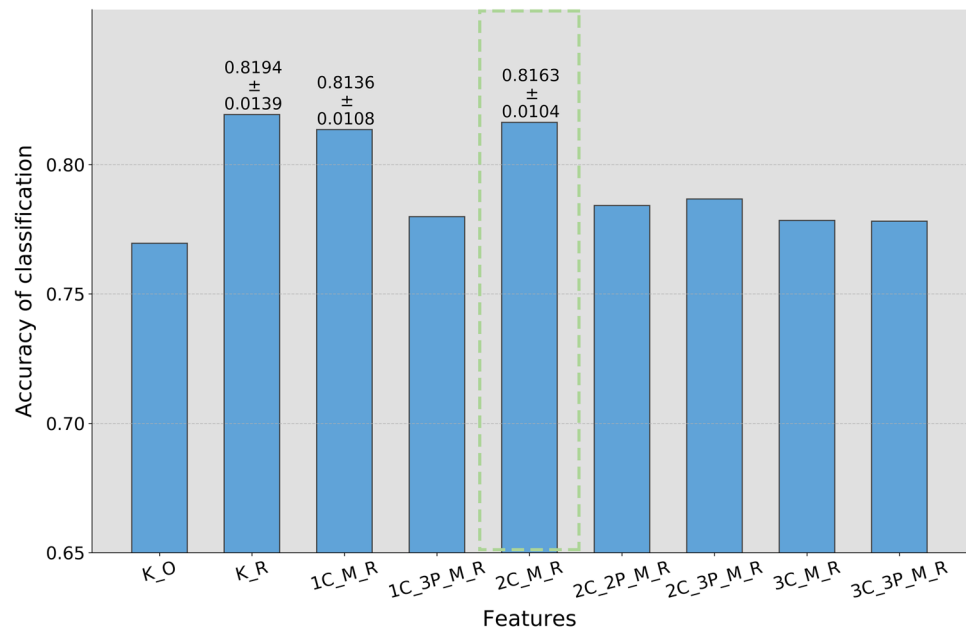
Fig. 10 The classification accuracy of SVM classifier based on features extracted by EEG-based copula model

are consistent with the results of Chen et al. The channel Oz highly related to visual stimuli is taken as an example. The correlation value between channels PO3 and Oz in RCM presents mild increases, which follows the findings in the study of Chen et al. The correlation relationship between channels C5 and Oz, and channels C6 and Oz in RCM both remain relatively stable. The correlation value between channels F3 and Oz in RCM shows significant changes, which may be explained by the fact that the frontal region has a connection to the visual process [47]. Therefore, the differences between the correlation matrices indicate that the RCM contains less pseudo-correlation.

Experiments on Assessing Visual Discomfort

To find the EEG features extracted by our proposed model highly related to visual discomfort, various linear concatenations of the smoothing parameters from MKDE

and the RCM are implemented. In Fig. 10, the horizontal coordinate presents different linear concatenations. The details of these extracted features are shown in Table 1. In Fig. 10, the top three highest accuracies of the SVM classifiers on different features are marked. The standard deviations of these three accuracies are also shown to verify the stability. The leftmost rectangle represents the accuracy of the SVM classifier based on feature K_O. The remaining rectangles present the accuracies of SVM on the features extracted by the proposed model. As Fig. 10 shows, the features extracted by the proposed model perform better. In other words, no matter how concatenated, these features have better discrimination to visual discomfort. The differences between the accuracies of the SVM classifiers on feature K_O and feature K_R indicate that RCM could greatly improve the discrimination of features to visual discomfort. As the differences in the remaining rectangles depict, although the number of smoothing

Table 1 The details of various extracted time-domain EEG features

| Features	Linear concatenation
K_O	Smoothing parameter from the KDE and the correlation matrix parameter of the copula function obtained by the original data
K_R	Smoothing parameter from the KDE and the RCM
1C_M_R	Optimal smoothing parameter from one cluster with the largest numbers of data and the RCM
1C_3P_M_R	First three smoothing parameters from one cluster with the largest numbers of data and the RCM
2C_M_R	Optimal smoothing parameters from two clusters with the first two largest numbers of data and the RCM
2C_2P_M_R	First two optimal smoothing parameters from two clusters with the first two largest numbers of data and the RCM
2C_3P_M_R	Three smoothing parameters from two clusters with the first two largest numbers of data and the RCM
3C_M_R	Optimal smoothing parameters from three clusters of data and the RCM
3C_3P_M_R	All the smoothing parameters from three clusters of data and the RCM
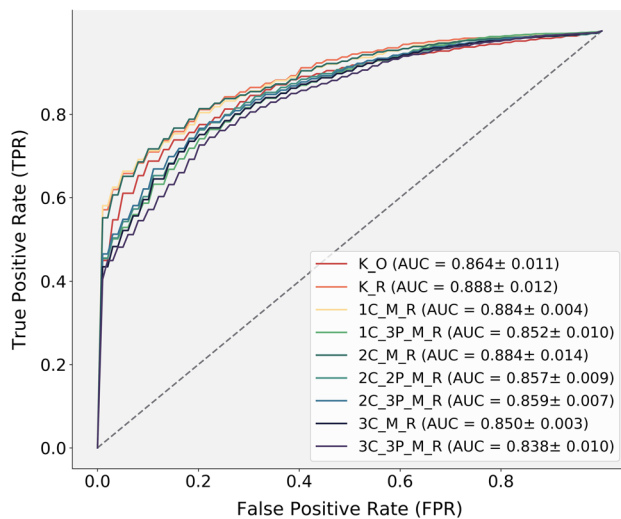
Fig. 11 The ROC curves of extracted features

parameters may contribute little to improving the accuracy of the classifier, it can keep the classifier stable. And the increased number of smoothing parameters may not always contribute to the accuracy of the SVM classifier. This may be explained by that the smoothing parameters corresponding to the cluster with the smallest data may be disturbed by redundant or irrelevant information. The SVM classifier based on feature K_R presents the highest accuracy among the three marked features, whereas the standard deviation is the largest. The feature 2C_M_R is considered the most promising feature to balance the trade-off between accuracy and stability, which has high discrimination to visual discomfort. The accuracy of the SVM classifier based on the feature 2C_M_R is 81.63%.

The receiver operating characteristic curve (ROC curve) to evaluate the extracted features further is shown in Fig. 11. The area under curve (AUC) value of feature K_O or features extracted by the proposed model is larger than 0.8. It means that features extracted by the

EEG-based copula model have high discrimination to visual discomfort. Compared to the remaining lines, the discrimination of features represented by orange, yellow, and dark green lines for visual discomfort is more precise, consistent with the results shown in Fig. 10. Therefore, the proposed model can extract the time-domain EEG features closely related to visual discomfort.

To testify the performance of the feature 2C_M_R extracted by the proposed model, the multiscale entropy (the state-of-the-art method) and sample entropy to assess discomfort are introduced. The AR model is also adopted to extract time-domain EEG features to compare with the feature 2C_M_R. Although the AR model has not been introduced into assessing visual discomfort, it is widely used to extract time-domain EEG features in other tasks, such as evaluating mental load [43] and classifying seizure [44]. Additionally, the statistics are performed on the features extracted by the AR model (hereinafter rough EEG features), such as the mean of rough EEG features and the standard deviation of rough EEG features, to explore the practical usefulness of the feature 2C_M_R. The results of the comparisons are shown in Table 2. The first column shows the method to extract the features. The second column presents the concrete features. The last column indicates the mean and standard deviation of the accuracy of the SVM classifiers based on extracted different features. The rough EEG features are not sensitive to visual discomfort. And the statistical features of rough EEG features do not show better discrimination to visual discomfort. All features extracted by entropy measures are less sensitive to visual discomfort compared with statistical features of rough EEG features. The features extracted by the proposed model perform better in discriminating visual discomfort despite the larger standard deviation. However, the value of the standard deviation is still much smaller than the mean value. Thereby, the features extracted by the proposed model present acceptable stability.

Table 2 The comparison of classification results based on the features extracted by different methods. *mean* and *SD* are the mean and standard deviation of the accuracy

Method of feature extraction	Features	Accuracy (*mean* $\pm$ *SD*)
AR model	coefficients (rough EEG features)	$0.4856 \pm \mathbf{0.0018}$
	median of coefficients	0.5060 ± 0.0075
	mean of coefficients	0.4947 ± 0.0065
	standard deviation of coefficients	0.5049 ± 0.0082
	kurtosis of coefficients	0.5098 ± 0.0055
	skewness of coefficients	0.5073 ± 0.0055
	combined statistics of coefficients	0.5059 ± 0.0024
Entropy measures	sample entropy	0.4873 ± 0.0863
	multiscale entropy (state-of-the-art method)	0.4562 ± 0.0945
Proposed model	2C_M_R	$\mathbf{0.8163} \pm 0.0108$

Discussion

Parameter Selection for MKDE

For MKDE, three parameters are needed to tune: the number of clusters C, the number of smoothing parameters m, and the coefficient α_j. Given the candidates of C, the GridSearch algorithm can exhaust all the clustering results of K-means. The C is selected according to the smallest within-cluster sum of squares. Given ranges of the number of smoothing parameters m, the GridSearch algorithm can exhaust all the results of estimated marginal density function with smoothing parameters. The m is selected according to the top several smallest mean error between the samples from the estimated marginal density function and the real data. Two strategies are used to find the coefficient α_j. The first strategy is to set the coefficient α_j equal the reciprocal of the number of smoothing parameters m. The second strategy is to determine the coefficient α_j is determined by the importance of each smoothing parameter to the estimated marginal density function. The importance of the smoothing parameters is determined by the reverse order of the sequence of error ratio, which is obtained by the ratio of the error of the estimated marginal density function by each smoothing parameter to the summation of the errors of the estimated marginal density function by all smoothing parameters. The coefficient α_j of each smoothing parameter equals to its importance value. According to a simulation experiment, there is no significant difference between the estimated marginal density functions based on the two strategies. To reduce the computation cost, the coefficient α_j is set to be the reciprocal of the number of smoothing parameters m.

Comparisons of Extracted Time-Domain EEG Features by Different Methods

In terms of Table 2, the time-domain EEG features extracted by the proposed method perform better than the features extracted by other methods, especially by the multiscale entropy (state-of-the-art method) proposed by Peng et al. Unlike the stereoscopic stimulus to induce the visual discomfort in this paper, Peng et al. designed the 2D stimuli with different flicking frequencies to cause visual discomfort. The differences between the designed stimuli may lead to different brain activities, which could explain that the features extracted by the multiscale entropy are less sensitive to visual discomfort induced by stereoscopic displays than the features extracted by the proposed model.

Conclusions

In this paper, the MKDE and RCM on the EEG-based copula model are proposed to improve the current performance of the EEG-based copula model and extract time-domain EEG features to assess visual discomfort better. As the simulation results present, the MKDE combined by the KDE with the K-means and the multiscale strategy can explore the fine-grained information in the data and better estimate the marginal density function. The RCM estimated by correlation analysis based on the remedied data obtained through the locality-sensitive dictionary learning can be closer to the real correlation matrix parameter of the copula density function. With the verified quality, the proposed method is used to extract time-domain EEG features, which are supposed to present high discrimination to visual discomfort for better assessing the visual discomfort. Compared with features extracted by other methods, the extracted features by the proposed method can be the more promising indicator for assessing visual discomfort. These findings may encourage more researchers to introduce the copula model into the analysis of EEG signals.

Although some progress has been made, this paper still has some limitations. Compared with KDE, current MKDE provides a less precise estimation of the marginal density function of heavily contaminated data compared to KDE. Additionally, waveforms of the time-domain diagram of coefficients of the AR model presents regularity with time. Therefore, the regularity of the waveforms can be considered during clustering instead of just the magnitude of the amplitudes. To reduce the computation cost, the Gaussian copula density function is introduced for its wide applications on time series, especially the financial time series. However, few investigations have shown which kinds of copula density functions are suitable to analyse EEG signals. Thereby, the types of copula density functions may contribute to the extraction of the time-domain EEG features. All the limitations encourage us to study further the application of the copula model on the analysis of EEG signals.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in this study involving human participants were in accordance with the ethical standards of the *Beijing Normal University, Beijing, China*.

References

1. Pölönen M, Järvenpää T, Bilcu B. Stereoscopic 3D entertainment and its effect on viewing comfort: Comparison of children and adults. Appl Ergon. 2013;44(1):151–60.

2. Gadia D, Granato M, Maggiorini D, Ripamonti LA, Vismara C. Consumer-oriented Head Mounted Displays: analysis and evaluation of stereoscopic characteristics and user preferences. Mobile Netw and Appl. 2018;23(1):136–46.

3. Henn JS, Lemole GM, Ferreira MA, Gonzalez LF, Schornak M, Preul MC, et al. Interactive stereoscopic virtual reality: a new tool for neurosurgical education. J Neurosurg. 2002;96(1):144–9.

4. Cui D, Wilson TD, Rockhold RW, Lehman MN, Lynch JC. Evaluation of the effectiveness of 3D vascular stereoscopic models in anatomy instruction for first year medical students. Anat Sci Educ. 2017;10(1):34–45.

5. Zou B, Liu Y, Guo M, Wang Y. EEG-based assessment of stereoscopic 3D visual fatigue caused by vergence-accommodation conflict. J Disp Technol. 2015;11(12):1076–83.

6. Park MC, Mun S. Overview of measurement methods for factors affecting the human visual system in 3D displays. J Disp Technol. 2015;11(11):877–88.

7. Lambooij M, Fortuin M, Heynderickx I, IJsselsteijn W. Visual discomfort and visual fatigue of stereoscopic displays: A review. J Imaging Sci Technol. 2009;53(3):30201–1.

8. Wang Y, Liu T, Li S, Wang J. Using electroencephalogram spectral components to assess visual fatigue caused by sustained prism-induced diplopia. In, 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). IEEE. 2016;2016:1551–6.

9. Monteiro TG, Skourup C, Zhang H. Using EEG for Mental Fatigue Assessment: A Comprehensive Look Into the Current State of the Art. IEEE Trans Hum Mach Syst. 2019;49(6):599–610.

10. Chen C, Li K, Wu Q, Wang H, Qian Z, Sudlow G. EEG-based detection and evaluation of fatigue caused by watching 3DTV. Displays. 2013;34(2):81–8.

11. Chen C, Wang J, Li K, Wu Q, Wang H, Qian Z, et al. Assessment visual fatigue of watching 3DTV using EEG power spectral parameters. Displays. 2014;35(5):266–72.

12. Kang MK, Cho H, Park HM, Jun SC, Yoon KJ. A wellness platform for stereoscopic 3D video systems using EEG-based visual discomfort evaluation technology. Appl Ergon. 2017;62:158–67.

13. Cho H, Kang MK, Yoon KJ, Jun SC. Feasibility study for visual discomfort assessment on stereo images using EEG. In: 2012 International Conference on 3D Imaging (IC3D). IEEE; 2012. p. 1–6.

14. Cho H, Kang MK, Ahn S, Kwon M, Yoon KJ, Kim K, et al. Cortical responses and shape complexity of stereoscopic image-A simultaneous EEG/MEG study. Neurosignals. 2016;24(1):102–12.

15. Peng Y, Wong CM, Wang Z, Wan F, Vai MI, Mak PU, et al. Fatigue Evaluation Using Multi-Scale Entropy of EEG in SSVEP-Based BCI. IEEE Access. 2019;7:108200–10.

16. Letham B, Sun W, Sheopuri A. Latent variable copula inference for bundle pricing from retail transaction data. In: International Conference on Machine Learning; 2014. p. 217–225.

17. Fan X, Xu RYD, Cao L. Copula mixed-membership stochastic block model. In: IJCAI International Joint Conference on Artificial Intelligence; 2016. .

18. Xu J, Wei W, Cao L. Copula-based high dimensional cross-market dependence modeling. In: 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA). IEEE; 2017. p. 734–743.

19. Nguyen-Huy T, Deo RC, Mushtaq S, Kath J, Khan S. Copula statistical models for analyzing stochastic dependencies of systemic drought risk and potential adaptation strategies. Stoch Environ Res Risk Assess. 2019;33(3):779–99.

20. Iyengar SG, Dauwels J, Varshney PK, Cichocki A. Quantifying EEG synchrony using copulas. In: 2010 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE; 2010. p. 505–508.

21. Qian D, Wang B, Qing X, Zhang T, Zhang Y, Wang X, et al. Drowsiness detection by Bayesian-copula discriminant classifier based on EEG signals during daytime short nap. IEEE Transactions on Biomedical Engineering. 2016;64(4):743–54.

22. Fontaine C, Frostig RD, Ombao H. Modeling non-linear spectral domain dependence using copulas with applications to rat local field potentials. Econom Stat. 2020;15:85–103.

23. Sklar M. Fonctions de repartition an dimensions et leurs marges. Publ inst statist univ Paris. 1959;8:229–31.

24. Scott DW. Multivariate density estimation: theory, practice, and visualization. John Wiley & Sons; 2015.

25. Zhang W, Zhang Z, Chao HC, Tseng FH. Kernel mixture model for probability density estimation in Bayesian classifiers. Data Min Knowl Discov. 2018;32(3):675–707.

26. Holmström L. The accuracy and the computational complexity of a multivariate binned kernel density estimator. J Multivar Anal. 2000;72(2):264–309.

27. Schwander O, Nielsen F. Learning mixtures by simplifying kernel density estimators. In: Matrix Information Geometry. Springer; 2013. p. 403–426.

28. Girolami M, He C. Probability density estimation from optimally condensed data samples. IEEE Transactions on pattern analysis and machine intelligence. 2003;25(10):1253–64.

29. Figueiredo MAT, Jain AK. Unsupervised learning of finite mixture models. IEEE Transactions on pattern analysis and machine intelligence. 2002;24(3):381–96.

30. Min A, Czado C. Bayesian inference for multivariate copulas using pair-copula constructions. J Financ Econ. 2010;8(4):511–46.

31. Brunel NB, Lapuyade-Lahorgue J, Pieczynski W. Modeling and unsupervised classification of multivariate hidden Markov chains with copulas. IEEE Transactions on Automatic Control. 2009;55(2):338–49.

32. Chicheportiche R, Chakraborti A. Copulas and time series with long-ranged dependencies. Phys Rev E. 2014;89(4):042117.

33. Kashefpoor M, Rabbani H, Barekatain M. Supervised dictionary learning of EEG signals for mild cognitive impairment diagnosis. Biomed Signal Process Control. 2019;53:101559.

34. Sharghian V, Rezaii TY, Farzamnia A, Tinati MA. Online Dictionary Learning for Sparse Representation-Based Classification of Motor Imagery EEG. In: 2019 27th Iranian Conference on Electrical Engineering (ICEE). IEEE; 2019. p. 1793–1797.

35. Yu K, Zhang T, Gong Y. Nonlinear learning using local coordinate coding. In: Advances in neural information processing systems; 2009. p. 2223–2231.

36. Wei CP, Chao YW, Yeh YR, Wang YCF. Locality-sensitive dictionary learning for sparse representation based classification. Pattern Recognition. 2013;46(5):1277–87.

37. Liu X, Song M, Tao D, Zhou X, Chen C, Bu J. Semi-supervised coupled dictionary learning for person re-identification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2014. p. 3550–3557.

38. Pang J, Zhang C, Qin L, Zhang W, Qing L, Huang Q, et al. Online dictionary learning for local coordinate coding with locality coding adaptors. Neurocomputing. 2015;157:61–9.

39. Pang J, Huang Q, Yin B, Qin L, Wang D. Theoretical analysis of learning local anchors for classification. In: Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012). IEEE; 2012. p. 1803–1806.

40. Hoffman DM, Girshick AR, Akeley K, Banks MS. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. J Vis. 2008;8(3):33–33.

41. Shibata T, Kim J, Hoffman DM, Banks MS. The zone of comfort: Predicting visual discomfort with stereo displays. J Vis. 2011;11(8):11–11.

42. Zheng Y, Zhao X, Yao L. The assessment of the visual discomfort caused by vergence-accommodation conflicts based on EEG. Journal of the Society for Information Display. 2019;27(10):608–18.

43. Nai-Jen H, Palaniappan R. Classification of mental tasks using fixed and adaptive autoregressive models of EEG signals. In: The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. vol. 1. IEEE; 2004. p. 507–510.

44. Subasi A, Alkan A, Koklukaya E, Kiymik MK. Wavelet neural network classification of EEG signals by using AR model with MLE preprocessing. Neural Netw. 2005;18(7):985–97.

45. Breiman L. Random forests. Machine learning. 2001;45(1):5–32.

46. Chen LC, Sandmann P, Thorne JD, Herrmann CS, Debener S. Association of concurrent fNIRS and EEG signatures in response to auditory and visual stimuli. Brain Topogr. 2015;28(5):710–25.

47. Roe AW, Parker AJ, Born RT, DeAngelis GC. Disparity channels in early vision. J Neurosci. 2007;27(44):11820–31.