IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 10, 2020, accepted July 18, 2020, date of publication July 24, 2020, date of current version August 19, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3011740

Research on Human-Computer Interaction Intention Recognition Based on EEG and Eye Movement

MINRUI ZHAO[1], HONGNI GAO[1], WEI WANG[1], AND JUE QU[1,2]
[1]Air and Missile Defense College, Air Force Engineering University, Xi'an 710051, China
[2]School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China

Corresponding authors: Jue Qu (qujue402@sina.com) and Minrui Zhao (zmr0204@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 51675530.

ABSTRACT In this work, we present a novel method to intention recognition, based on electroencephalogram (EEG) and eye movement in human-computer interaction(HCI). The fusion of EEG and eye movement will allow the utmost of the advantages of the two physiological signals. Signals of EEG and eye movement were collected for feature extraction, recognition network of machine learning pattern was input for intent recognition, final recognition result was attained by decision-level fusion. We compare the results of the Intention Recognition Algorithms to those of an experiment involving the intention recognition of the operator in a simulated flight mission. In most every case, results show that the intention recognition algorithms performed better than solely rely on single signal.

INDEX TERMS Intention recognition, physiological signals, EEG, eye movement, simulated flight, decision-level fusion.

I. INTRODUCTION

One of the key targets of the human-computer interaction intelligence [1] is to improve the user's intention perception derived from the human-computer interaction system. It provides theoretical basis and technical support for the design of the adaptive human-computer interaction system as well as reducing personal errors during the operation [2]. Current common human-computer interaction intention recognition methods mainly rely on signals of EEG or eye movement.

Brain-computer interface (BCI) is an application form of EEG signals, which can establish a communication relationship between the human brain and external devices, thus enabling the brain to control external devices [3]–[5]. BCI has been applied in fields like medicine [6], [7], neurobiology [8], psychology [9], [10]. Motor Imagery (MI) Electroencephalogram (EEG) boasts the characteristics of flexibility, non-invasiveness, low environmental requirements and high resolution. Therefore, MI is one of the widely application forms of BCI [11]. The spectrum power of EEG signal during motion imagination will vary with the content of MI task, which is called event-related synchronization/ desynchronization (ERS/ERD)[12]. Razi *et al.* [11] extract features of EEG signal and use the machine learning algorithms to classify it. The average Kappa value of motor imaging recognition is 75%. Amin *et al.* [13] collect EEG signals, perform feature extraction and data classification to realize the remote control of UAV through BCI. However, these intent recognition methods only rely on EEG signals without the advantages of integrating eye movement signals, of which recognition accuracy rate needs to be further improved.

Studies have shown that visual channels provide more than 80% of the external information to people. In recent years, many scholars have studied its user's intention of human-computer interaction based on human visual behavior. Deng M. *et al.* [14] used eye movement data to analyze the user's behavioral intention and emotional experience. Jang *et al.* [15] invented a visual search intention recognition method based on eye movement patterns and pupil characteristics. The results show that accuracy rate can reach more than 90% when layer support vector machine recognition is adapted. In addition, eye movement tracking is also widely applied in user interaction behavior analysis [16], user visual search analysis [17] and visual stimulus interest analysis [18] and other fields. These methods only rely on eye movement signals without fusing EEG signals and fail to make full use of EEG signals to analyze the cognitive state of human brain.

The associate editor coordinating the review of this manuscript and approving it for publication was Eunil Park.

In recent years, many scholars have tried to integrate multiple physiological information to improve the accuracy of human-computer interaction recognition. Park U *et al.* [19] integrated signals of EEG and eye movement to identify intentions, and eventually found that the accuracy of integrated signals of EEG and eye movement recognition is about 5% higher than that of relying on one single physiological signal. Postelnicu C *et al.* [20] combined eye movement, EEG and gesture characteristics to control the 6-degree-of-freedom manipulator. The results show that the SUS score is higher than average. Chowdhury *et al.* [21] summarized the detection methods of automobile drivers' fatigue and found that the accuracy of fatigue recognition based on multiple physiological signals was significantly higher than that based on one single physiological signal.

Accordingly, this paper has proposed a human-computer interaction intention recognition method based on the fusion of EEG and eye movement and the introduction of decision-layer fusion, which can perform intention recognition while the user is performing human-computer interaction. The specific implementation process is as follows: collecting EEG and eye movement signals of its user for feature extraction; using pattern recognition algorithms to classify and identify physiological signal features; performing decision-level fusion on the classifiable algorithm to obtain the final result, performing user intention-induced experiments to verify the feasibility of the method. The effect of different EEG feature extraction methods and the effect of different machine learning algorithms on recognition accuracy have also been compared.

II. METHODS

A. EEG FEATURE EXTRACTION BASED ON COMMON SPATIAL DOMAIN MODE(CSP)

Algorithm of CSP mode has proved to be effective in the analysis of EEG signals based on EDS/ERS, however the CSP mode algorithm is proposed for the binary classification problem. Therefore, the CSP algorithm needs to be improved when multi-classification problem arise, and compare any two types of categories in order. CSP finds the best projection direction by spatially projecting the original signal, which satisfies that the first type signal has the largest variance in a certain direction and the second type signal has the smallest variance, thus the projection direction with the largest difference between the two types is obtained.

E is set as the EEG signal matrix after removing the DC component, the dimension of which is $N \times T$. Where N is the number of EEG data channels, T is the number of sampling points for each channel. Therefore, the covariance matrix of the EEG data is,

$$C_i = \frac{EE^{\mathrm{T}}}{\mathrm{trace}\left(EE^{\mathrm{T}}\right)} \tag{1}$$

where E^{T} is the transposed matrix of E, $trace\left(EE^{\mathrm{T}}\right)$ is the trace of EE^{T}. The average covariance of various EEG signals is $C_i, i \in \{1, 2\}$. Therefore, the sum of the average covariance matrix is:

$$C_c = C_1 + C_2 \tag{2}$$

When C_c is eigenvalue decomposd, it can be known:

$$C_c = U_c \lambda_c U_c^T \tag{3}$$

where U_c is the eigenvector matrix of C_c, λ_c is the eigenvalue matrix of C_c,.When whitening matrix $P = \sqrt{\lambda_c^{-1}} U_c^{\mathrm{T}}$ is constructed, it can be known:

$$I = PC_cP^{\mathrm{T}} \tag{4}$$

where I is the identity matrix. When C_1, C_2 is transformed, it can be known:

$$S_i = PC_iP^{\mathrm{T}}, \quad i \in \{1, 2\} \tag{5}$$

where, S_1 and S_2 are of the same eigenvector, of which the corresponding eigenvalue sum is 1. In other words, S_2 is of the minimal eigenvalue when S_1 is in the direction of the maximal eigenvalue; in the direction of the smallest eigenvalue, S_2 is of the maximal eigenvalue when S_1 is in the direction of the minimal eigenvalue otherwise.

If

$$S_1 = B\lambda_1 B^{\mathrm{T}} \tag{6}$$

It can be known:

$$\begin{cases} S_2 = B\lambda_2 B^{\mathrm{T}} \\ \lambda_1 + \lambda_2 = \mathbf{I} \end{cases} \tag{7}$$

The projection matrix is

$$W = B^{\mathrm{T}} P \tag{8}$$

It can be seen that W is matrix of $\mathrm{N} \times \mathrm{N}$ order, the original signal E is projected to obtain a new signal:

$$Z = WE \tag{9}$$

m is generated by projected the selected first m rows and the last rows. New signal is transformed as follow to obtain the final eigenvalue:

$$f_j = \lg \left[\frac{var\left(\mathbf{Z}_j\right)}{\sum_{k=1}^{2m} var\left(\mathbf{Z}_k\right)} \right], \quad j = 1, 2 \cdots 2m \tag{10}$$

where $var\left(\mathbf{Z}_j\right)$ is the variance of $\mathbf{Z}_j$.

For any four types of signals, every two types of them are processed by CSP, and 6 projection matrices can be obtained. For each matrix W, 4 optimal directions are selected front and back, thus 8 optimal directions are obtained, so is a $6 \times 8 = 48$ dimensional feature vector. Then use discriminant method of Fisher to reduce the dimension, which can reduce the 48-dimension feature vector to a 3-dimension feature vector. Finally, the Bayesian classifier is used to classify the features after dimensionality reduction according to the prior probability theory:

$$P(y_i|f) = \frac{P(f|y_i)\,P(y_i)}{\sum_{j=1}^{C} P(f|y_j)\,P(y_j)} \tag{11}$$

where $P(y_i)$ is the a priori probability of data labeled type i, $P(y_i|f)$ is the posterior probability of the sample f whose characteristics belong to type i, $P(f|y_i)$ is the likelihood ratio of feature f generated by type i. The classification result is i of the highest value of posterior probability.

B. EXTRACTION OF EYE MOVEMENT SIGNAL FEATURE

RED5 eye tracker is applied to collect eye movement data with sampling frequency reaching 500 Hz. The eye movement feature can reflect the user's eye movement behavior. In this paper, five physiological characteristics are collected for analysis, they are of fixation point X coordinate (FX), fixation point Y coordinate (FY), pupil diameter (PD), fixation time (FT) and saccade amplitude (SA). The five types of eye movement features mentioned above can be directly obtained from the data analysis software of the eye tracker.

The fixation point X coordinate (FX) and fixation point Y coordinate (FY) represent the direction of fixation point X and fixation point Y of user during his or h human-computer interaction, which can reflect the position of user's fixation point on the screen.

Pupil diameter (PD) can be regarded as one of the indicators reflecting the user's real-time cognitive load. When the cognitive load increases, the pupil diameter increases; the pupil diameter will decrease otherwise. Therefore, user's cognitive state can be reflected through the indicator.

The fixation time (FT) can reflect the difficulty of its user in processing visual information. The processing time will be longer when faced with complex information, so this indicator can reflect the user's cognitive state.

Usually, the saccade amplitude (SA) is reflected as the amplitude between fixation points, reflecting the difficulty of user for processing the visual information. When visual information is roughly processed, the saccade amplitude becomes larger; Otherwise, when the user When visual information is smoothly processed, the saccade amplitude turns smaller. Therefore, it can reflect the user's cognitive state.

C. SVM CLASSIFIER

SVM classification algorithm is a machine learning classification algorithm based on statistical learning theory. It is different from the ordinary optimization algorithm in pursuit of minimum experience risk. SVM algorithm improves the generalization ability of the algorithm, minimizes experience risk and confidence range by minimizing the structured risk, preferably solving the problems of over-learning, model selection, dimensionality disaster and nonlinearity in algorithm of pattern recognition under small sample conditions. The essential of the algorithm is to find the optimal classification plane that maximizes the classification interval between the two types.

Let the sample set be:

$$D = \{(x_1, y_1), \cdots, (x_l, y_l)\} \quad x \in \mathbf{R}^n \\ y \in \{+1, -1\} \tag{12}$$

where n is sample dimension, l is the number of samples, the classification plane hyperplane of n dimension, which can be expressed as:

$$\langle w, x\rangle + b = 0 \quad w \in \mathbf{R}^n\ b \in \mathbf{R} \tag{13}$$

Thus the classification interval can be $2/\|w\|$. Obviously, when $\|w\|$ takes the minimum value, the classification interval reaches the maximum. Therefore, the question to obtain the maximum classification interval can be transformed into solving the following constrained optimization problem:

$$\begin{aligned} &\min\ \|w\|^2/2 \\ &\text{s.t. } y_i(w \cdot x_i + b) - 1 \ldots 0, \quad i = 1, 2, \cdots, l \end{aligned} \tag{14}$$

The solution vector w^* is the support vector when the problem is linearly separable, and the classification function of the support vector machine is determined by the support vector. Support vector machine describes the decision function of the optimal classification hyperplane, which can be expressed as:

$$f(x) = \operatorname{sgn}\left(\langle w^*, x\rangle + b\right) \tag{15}$$

The method of projecting the non-linear transformation of the sample into a high-dimensional space is usually used to make the sample separable in this high-dimensional feature space when the problem is linearly inseparable. According to the Mercer condition, the decision function is:

$$f(x) = \operatorname{sgn}\left(\sum_{i=1}^{l} \alpha_i y_i K(x, x_i) + b\right) \tag{16}$$

where $K(x, x_i)$ is the kernel function, α_i is Lagrange multiplier.

D. THEORY OF D-S EVIDENCE

Theory of D-S evidence is an imprecise reasoning method proposed by Dempster and perfected by Shafer, his student. It can be applied to deal with uncertain information, and the required conditions are less strict than theory of Bayesian probability. The theory can not only deal with the uncertainty caused by imprecise prior knowledge, but it also can handle the uncertainty caused by ignorance.

Set Θ as recognition framework, m is a credibility distribution function on the interval [0, 1], also known as the Mass function, indicating the happening degree to the evidence supports event A. $m(\emptyset) = 0$, and it satisfies:

$$\sum_{A \subseteq \Theta} m(A) = 1 \tag{17}$$

For $\forall\{A, B, C\} \subseteq \Theta$, according to the Dempster synthesis rule, any two Mass functions of Θ, the synthesis method

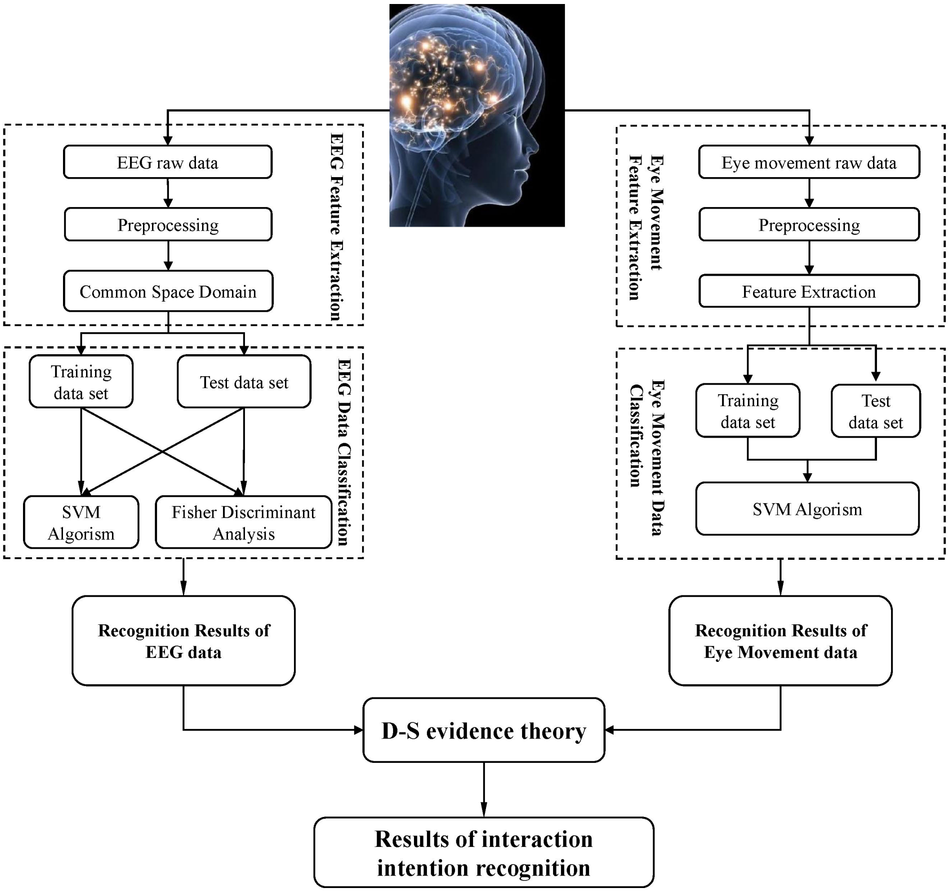


FIGURE 1. Diagram of Physiological Signals-based Intention Recognition.

of m_1 and m_2 is:

$$m_1 \oplus m_2(A) = \frac{1}{K} \sum_{B \cap C = A} m_1(B) \cdot m_2(C) \tag{18}$$

$$K = \sum_{B \cap C \neq \emptyset} m_1(B) \cdot m_2(C) \tag{19}$$

For $\forall A \subseteq \Theta$, according to the Dempster synthesis rule, the synthesis method of any finite Mass function $m_1, m_2, \cdots, m_n$ of Θ is

$$\begin{aligned} m(A) &= (m_1 \oplus m_2 \oplus \cdots \oplus m_n)(A) \\ &= \frac{1}{K} \sum_{A_1 \cap A_2 \cap \cdots \cap A_n = A} m_1(A_1) \cdot m_2(A_2) \cdots m_n(A_n) \end{aligned} \tag{20}$$

$$K = \sum_{A_1 \cap A_2 \cdots \cap A_n \neq \emptyset} m_1(A_1) \cdot m_2(A_2) \cdots m_n(A_n) \tag{21}$$

E. HUMAN-COMPUTER INTERACTION INTENTION RECOGNITION OF DECISION-LEVEL FUSION

The process of decision-level fusion is shown in the Figure 1. First, pre-processing and feature extraction are performed on the collected physiological signals to obtain feature vectors corresponding to the physiological signals, whereafter the EEG and eye movement signal feature vectors are respectively classified by a classifier, and finally the classification results of each classifier are fused under DS evidence theory to obtain the fusion result of decision layer.

Assuming that $A_1, A_2, \cdots, A_k$ is cognitive intent of k types classified by quasi-physiological signals, the recognition framework is

$$\Theta = \{A_1, A_2, \cdots, A_k\} \tag{22}$$

The dentification algorithm of recognition function of each physiological signal feature for each type of intent credibility is:

$$m_i(A_1, A_2, \cdots, A_k, \Theta) = (p_i q_{i1}, p_i q_{i2}, \cdots, p_i q_k, 1 - p_i) \tag{23}$$

where m_i is the credibility distribution function of physiological feature recognition algorithm of type i, $i = 1, 2 \cdots n$.

The correct rate of the first physiological feature recognition algorithm

Where p_i is the credibility of the physiological characteristic identification of type i. q_{ij} is the judgment sample of the physiological characteristic identification of type i as the credibility of cognitive intention of type j, $j = 1, 2 \cdots k$.

For any cognitive intent A_j in the recognition framework Θ, the rule of Dempster decision rule adapted multiple physiological feature classification is:

$$m(A_j) = (m_1 \oplus m_2 \oplus \cdots \oplus m_n)(A_j) \quad j = 1, 2, \cdots, k \tag{24}$$

FIGURE 2. The Experimental environment.

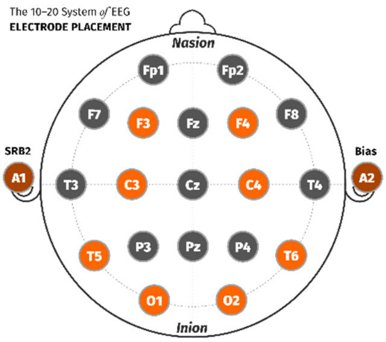


FIGURE 3. Location of electrode.

It can be obtained from equation (23), equation (24) and equation (19)

$$m\left(A_j\right) = \frac{1}{K}\left[\sum_{i=1,2,\ldots,n}(1-p_i)\prod_{h=1,2,\cdots,n;h\neq i}p_h q_j + \prod_{i=1,2,\ldots,n}p_i q_j\right] \tag{25}$$

$$K = \sum_{i=1,2,\cdots,n}(1-p_i)\prod_{h=1,2,\cdots,n\neq i}p_h q_j + \prod_{i=1,2,\cdots,n}p_h q_j + \prod_{i=1,2,\cdots,n}(1-p_i) \tag{26}$$

where the algorithm accuracy rate p_i is usually is determined by the correct accuracy rate of training sample, and the credibility q_{ij} of cognitive intention is determined by the algorithm output calculation:

$$q_{ij} = \frac{1}{2C_k^2}\left[\sum_{j\neq h}F_i\left(A_j, A_h\right) + k - 1\right] \quad j, h = 1, 2, \cdots, k \tag{27}$$

where j, h is cognitive intention serial number, k is the number of cognitive intention types, $F_i\left(A_j, A_h\right)$ is the judgment based on cognitive intention and classification, $F_i\left(A_j, A_h\right) \in \{+1, -1\}$.

Not only the category of unknown samples but also the probability that the samples belong to each category can be output when D-S decision-level fusion interaction intention discriminating method is used. For the D-S evidence theory, the credibility allocation and assignment problem is effectively and intuitively solved by the overall classification accuracy rate obtained by machine learning algorithm training.

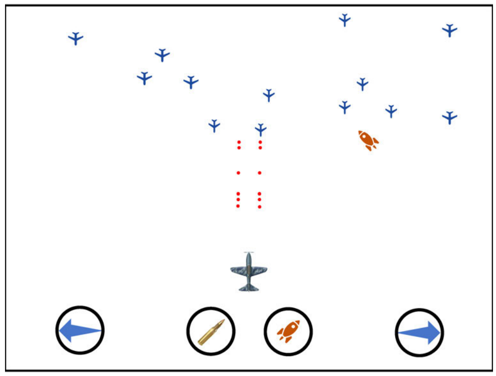


FIGURE 4. Experimental interface of motor imaginary.

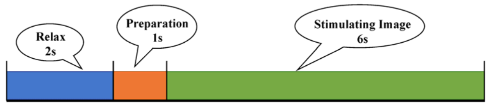


FIGURE 5. Experimental paradigm of motor imaginary.

III. EXPERIMENTS

A. PARTICIPANTS

In order to verify the scientificity and effectiveness of the proposed human-computer interaction intention recognition method, it is necessary to collect EEG and eye movement data of the user during his or her human-computer interaction. 24 male users (four of whom could not be used as effective participants because the data collection rate was less than 50%), with an age range of 18 to 22 years (M = 23.4, SD = 2.1) were recruited in this research. Before the experiment, each user was familiar with the experiment process and precautions and had signed an informed consent form.

B. EXPERIMENT PRODUCERS

The experiment environment is as shown in the Figure 2. The curtains were closed during the experiment. Only experiment users and operators could enter and close other electronic devices to eliminate light changes and other electromagnetic signal interference during each experiment. Experimental equipment consists of DELL computer, RED5 eye movement tester and Neuroscan 32 brain conduction electroencephalograph. The sampling frequency of the eye movement signal was 50 Hz; the display resolution was 1280 × 1024 pixels, and the screen brightness was 300 cd/m2; the distance between the user and the screen was about 60 cm, and the eyes of the user are about the same height as the center of the screen. The electrode distribution of the Neuroscan 32 brain conduction electroencephalograph adopted international standard 10-20. The left mastoid was

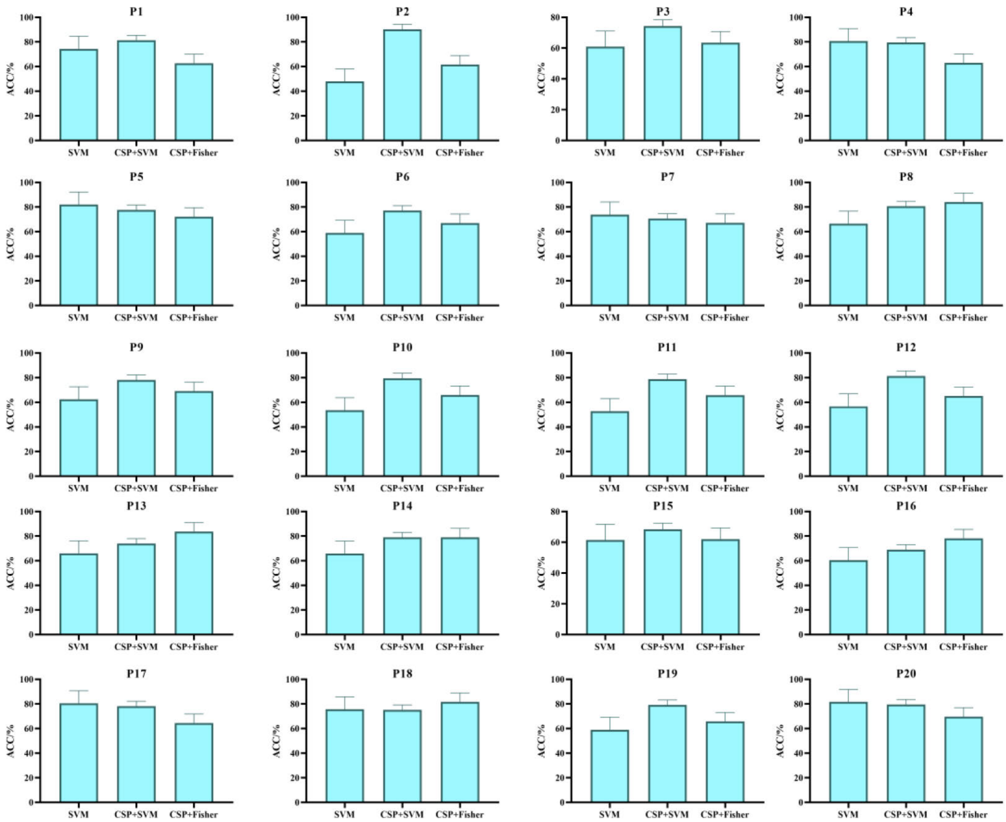


FIGURE 6. **Recognition accuracy under different processing methods.**

TABLE 1. **Accuracy of eye movement classification prediction under different feature.**

Index of Eye Movement	Recognition Rate/%
pupil diameter & fixation time	37.15
pupil diameter & saccade amplitude	50.13
gaze time & saccade amplitude	44.97
(FX&FY)& pupil diameter	63.85
(FX&FY)& saccade amplitude	69.88
(FX&FY)& gaze time	78.13
(FX&FY)& pupil diameter & fixation time	78.15
(FX&FY)& pupil diameter & saccade amplitude	81.56
(FX&FY)& fixation time & saccade amplitude	80.64
pupil diameter & fixation time & saccade amplitude	54.93
(FX&FY) & pupil diameter & fixation time & saccade amplitude	85.34

used as the reference electrode, and the middle prefrontal lobe was the ground electrode. Additionally, the vertical and horizontal channel electrooculogram signals were collected at a sampling frequency of 250 Hz. 50 Hz notch and 0.05 $\sim$ 10 Hz online band-pass filter was adapted to ensure that the electrode impedance is less than 5kΩ, EMG and electrooculogram artifacts were removed after the signal is collected.

TABLE 2. **Average classification prediction accuracy under different EEG processing algorithm.**

Processing method	SVM	CSP+SVM	CSP+Fisher
Average recognition accuracy	66.09	77.08*	69.59
Standard deviation	10.20	3.83	7.32

The classic motor imagination experiment paradigm was adopted in the research. As shown in the figure 5, the screen displayed "relax" before the operation imagination starts for 2s, and the user was relaxed and ready to start; then the screen displayed "preparation" for 1s, prompting the user to prepare to start operation imagination. Next, the screen presented the operation interface. The user's operations included "Move left", "Move right", "Attack" and "Launch a missile". The icon turned yellow when the user needs to operate the imagination, and the user performed different operation imaginations according to the different prompts. The total time for each operation is 9s.

C. RESULTS

In the experiment of 3.1, we collected five eye movement parameters: fixation point X coordinate (FX), fixation point Y coordinate (FY), pupil diameter (PD), fixation time (FT), and saccade amplitude (SA). FX and FY synthesize a type of feature as a fixation point feature. The extracted four types of eye movement features of the test user were used as the basis for algorithm classification, and the eye movement data

TABLE 3. Classification accuracy with EEG & eye movement data in training set.

Subject Number	Recognition Rate/%	
	EEG	Eye Movement
P1	79.63	87.29
P2	79.67	85.55
P3	68.70	87.82
P4	71.15	84.84
P5	79.81	81.26
P6	79.20	81.60
P7	74.03	88.15
P8	77.65	85.02
P9	72.68	84.44
P10	71.52	86.69
P11	71.00	81.10
P12	77.90	88.33
P13	73.02	83.65
P14	69.29	83.91
P15	76.42	78.65
P16	68.77	80.15
P17	76.15	93.90
P18	74.76	83.38
P19	68.62	92.85
P20	77.53	84.93
Average	74.38	85.17

of 20 participants was preprocessed and had been performed feature extraction, 60% of which was used as the training set, and 40% of which was used as the test set. The classification accuracy was shown in the Table 1 when SVM algorithm was adapted to perform operation imagination classification. It can be seen from the Table 1 that the greater the number of eye movement indicators, the higher the recognition accuracy is, and the location characteristics of the fixation point have a greater impact on the accuracy, showing that the location feature of the fixation point can better reflect the user's intention, but since the fixation point was unstable in the experiment, it is necessary to combine other eye movement features to improve the recognition reliability. The five eye movement features selected in this research had an accuracy rate of 85.34% for the cognitive intention experiment, indicating that the selection of eye movement indicators is effective.

Four methods, SVM, CSP+SVM, CSP+Fisher, were selected to explore the impact of different machine learning methods and data preprocessing methods on human-computer interaction intention recognition to identify the EEG data set of 20 participants, 60% was used as the training set and 40% was used as the test set. The final recognition accuracy rate of each subject is shown in Figure 6. It can be seen that for different participants, the accuracy rate of the algorithm recognition is slightly different from the standard deviation. The recognition results of the different EEG processing algorithm are shown in Table 2. The recognition accuracy of the CSP+SVM algorithm is significantly higher than that of the other two algorithms ($P<0.5$), and the standard deviation of this method is the smallest, indicating that the algorithm has low sensitivity and strong generalization ability, which is suitable for processing EEG data. From the perspective of recognition accuracy, the average recognition accuracy of CSP method is 77.08%, which is higher than the average accuracy of 66.09% without adapting feature extraction method. In the case of adapting CSP feature extraction method, the average recognition accuracy of the SVM algorithm is 77.08%, while the average recognition rate of the Fisher method is 69.59%. It shows that for the CSP feature extraction method, the recognition accuracy of the SVM algorithm is slightly higher than that of the Fisher algorithm.

CSP+SVM method was adapted to preprocess the EEG data before SVM algorithm was trained. The feature extraction was performed on the eye movement data, and the

TABLE 4. Comparison of % accuracies of classifiers based on different feature.

Subject Number	Physiological Recognition Accuracy /%		
	EEG	Eye Movement	EEG&Eye Movement
P1	75.41	85.91	92.02
P2	83.16	93.65	87.22
P3	72.63	83.97	85.99
P4	83.52	96.77	88.17
P5	74.67	91.59	93.27
P6	64.99	83.21	88.67
P7	78.79	92.12	88.48
P8	76.36	80.15	91.94
P9	71.72	95.60	91.95
P10	83.60	77.71	91.48
P11	70.14	70.70	93.54
P12	69.79	75.69	92.11
P13	77.23	82.70	85.80
P14	62.23	77.21	89.02
P15	82.24	83.32	89.61
P16	69.22	81.62	86.13
P17	61.65	92.95	89.02
P18	65.43	74.40	85.39
P19	70.28	94.91	87.39
P20	78.95	77.90	87.24
Average	73.60	84.60	89.22
± Standard deviation	± 6.78	± 7.73	± 2.56

fixation point position (FX&FY), pupil diameter (PD), fixation time (FT) and saccade amplitude (SA) were selected. Four types of eye movement features were trained with SVM algorithm, and the cross-contrast method were used to determine the SVM algorithm parameters. The recognition results of the training set are shown in Table 3.

According to D-S theory, three cognitive intentions in this experiment constitute a recognition framework $\Theta = \{$Move left, Move right, Attack, Lunch a missile$\}$, and the average value of the sample of 20 participants in the Table 3 is used as the recognition accuracy rate, thus it can be known that $\{p_1, p_2\} = \{0.852, 0.744\}$, the eye movement and EEG recognition methods are caused by the uncertainty resulted from ignorance $\{m_1(\Theta), m_2(\Theta)\} = \{0.148, 0.256\}$. The test samples corresponding to the training samples are separately classified by adapting the SVM algorithm and perform decision-level fusion classification based on the D-S theory. The recognition results are shown in Table 4.

The experiment results show that the recognition rate of the test sample is lower than that of the training sample, but the decrease is not sharp, the accuracy of EEG data decreases by 0.78%, and the accuracy of eye movement data recognition decreases by 0.57%, indicating that SVM algorithm is of strong generalization ability. Comparing the accuracy of eye movement and EEG data recognition, it is found that the accuracy of eye movement data recognition, 84.60%, is higher than that of EEG data recognition 73.60%, indicating that the accuracy of eye movement data for human-computer interaction intention recognition is higher than that of EEG data. According to DS evidence theory, the decision-level fusion of data can achieve a maximum recognition accuracy of 93.54%, and an average recognition accuracy of 89.22%, which is higher than the accuracy of eye movement and EEG data recognition, and the variance of the data recognition accuracy is only 2.56, which shows that the data fusion method based on DS theory has low sensitivity to samples and strong generalization ability, proving that D-S theory has advantages in the identification of multiple physiological information intentions.

IV. CONCLUSION

To solve the problem that the traditional human-computer interaction intention recognition accuracy is relatively low, and different physiological information cannot be effectively fused, EEG and eye movement information fusion human-computer interaction intention recognition method based on DS evidence theory was designed to recognize the user's human-computer interaction. intention. The EEG and eye movement signals were separately extracted and classified by collecting the user's raw data, whereafter the D-S evidence theory was adapted to fuse the EEG and eye movement signal classification results. The experiment results show that the EEG eye movement information fusion human-computer interaction intention recognition method based on D-S evidence theory has the characteristics of high accuracy and strong generalization ability, which lays the foundation for the future of adaptive design of human- interaction interface.

ACKNOWLEDGMENT

The authors would like to thank W. Yan for language support.

REFERENCES

[1] A.-V. L. María and M.-G. V. Rodrigo, "Enrichment of human-computer interaction in brain-computer interfaces via virtual environments," *Comput. Intell. Neurosci.*, vol. 2017, pp. 1–12, Jan. 2017.

[2] R. Aya, B. Mihaly, S. Piotr, G. Felix, S. Abdul, and V. Ivan, "Brain–computer interface spellers: A review," *Brain Sci.*, vol. 8, no. 4, pp. 57–64, 2018.

[3] P. J. Benson, "Decoding brain-computer interfaces," *Science*, vol. 360, no. 6389, pp. 615–616, 2018.

[4] K. Belwafi, O. Romain, S. Gannouni, F. Ghaffari, R. Djemal, and B. Ouni, "An embedded implementation based on adaptive filter bank for brain–computer interface systems," *J. Neurosci. Methods*, vol. 305, pp. 1–16, Jul. 2018.

[5] A. Chowdhury, R. Shankaran, M. Kavakli, and M. M. Haque, "Sensor applications and physiological features in drivers' drowsiness detection: A review," *IEEE Sensors J.*, vol. 18, no. 8, pp. 3055–3067, Apr. 2018.

[6] L. E. H. van Dokkum, T. Ward, and I. Laffont, "Brain computer interfaces for neurorehabilitation—Its current status as a rehabilitation strategy post-stroke," *Ann. Phys. Rehabil. Med.*, vol. 58, no. 1, pp. 3–8, Feb. 2015.

[7] M. Deng and X. Gu, "Information acquisition, emotion experience and behaviour intention during online shopping: An eye-tracking study," *Behav. Inf. Technol.*, vol. 2 pp. 1–11, Jan. 2020.

[8] S. Ge, R. Wang, Y. Leng, H. Wang, P. Lin, and K. Iramina, "A double-partial least-squares model for the detection of steady-state visual evoked potentials," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 4, pp. 897–903, Jul. 2017.

[9] M. Gil, M. Albert, J. Fons, and V. Pelechano, "Designing human-in-the-loop autonomous cyber-physical systems," *Int. J. Hum.-Comput. Stud.*, vol. 130, pp. 21–39, Oct. 2019.

[10] E. M. Hammer, S. Halder, S. C. Kleih, and A. Kübler, "Psychological predictors of visual and auditory P300 brain-computer interface performance," *Frontiers Neurosci.*, vol. 12, pp. 307–316, May 2018.

[11] J. Hautala, C. Kiili, Y. Kammerer, O. Loberg, S. Hokkanen, and P. H. T. Leppänen, "Sixth graders' evaluation strategies when reading Internet search results: An eye-tracking study," *Behav. Inf. Technol.*, vol. 37, no. 8, pp. 761–773, Aug. 2018.

[12] Y.-M. Jang, R. Mallipeddi, and M. Lee, "Identification of human implicit visual search intention based on eye movement and pupillary analysis," *User Model. User-Adapted Interact.*, vol. 24, no. 4, pp. 315–344, Oct. 2014.

[13] J. Kögel, J. R. Schmid, R. J. Jox, and O. Friedrich, "Using brain-computer interfaces: A scoping review of studies employing social research methods," *BMC Med. Ethics*, vol. 20, no. 1, pp. 1–17, Dec. 2019.

[14] R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof, and G. Pfurtscheller, "Brain–computer communication: Motivation, aim, and impact of exploring a virtual apartment," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 15, no. 4, pp. 473–482, Dec. 2007.

[15] J. Mou and D. Shin, "Effects of social popularity and time scarcity on online consumer behaviour regarding smart healthcare products: An eye-tracking approach," *Comput. Hum. Behav.*, vol. 78, pp. 74–89, Jan. 2018.

[16] A. Nourmohammadi, M. Jafari, and T. O. Zander, "A survey on unmanned aerial vehicle remote control using brain–computer interface," *IEEE Trans. Hum.-Mach. Syst.*, vol. 48, no. 4, pp. 337–348, Aug. 2018.

[17] C.-C. Postelnicu, F. Girbacia, G.-D. Voinea, and R. Boboc, "Towards hybrid multimodal brain computer interface for robotic arm command," in *Augmented Cognition* (Lecture Notes in Computer Science). Orlando, FL, USA: Springer, 2019, pp. 461–470.

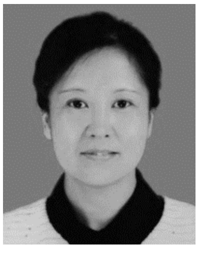
[18] Z. Qiu, B. Z. Allison, J. Jin, Y. Zhang, X. Wang, W. Li, and A. Cichocki, "Optimized motor imagery paradigm based on imagining chinese characters writing movement," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 7, pp. 1009–1017, Jul. 2017.

[19] U. Park, R. Mallipeddi, and M. Lee, "Human implicit intent discrimination using EEG and eye movement," *BMC Neurosci.*, vol. 17, no. 54, pp. 11–18, 2019.

[20] J. J. Shih, D. J. Krusienski, and J. R. Wolpaw, "Brain-computer interfaces in medicine," *Mayo Clinic Proc.*, vol. 87, no. 3, pp. 268–279, Mar. 2012.

[21] Y. Wang and B. A. Sparks, "An eye-tracking study of tourism photo stimuli: Image characteristics and ethnicity," *J. Travel Res.*, vol. 55, no. 5, pp. 588–602, May 2016.

MINRUI ZHAO received the B.S. degree from Air Force Engineering University, China, in 2018, where he is currently pursuing the M.S. degree. His research interests include human-computer interaction, human factor, ergonomics, and intelligent algorithm.



HONGNI GAO received the M.S. degree from Air Force Engineering University, China, in 2002. She is currently an Associate Professor with Air Force Engineering University. Her research interests include mechatronics design and human-computer interaction.

WEI WANG received the Ph.D. degree from Xi'an Jiaotong University, in 2008. He is currently a Professor with Air Force Engineering University. His research interests include mechatronics design, artificial intelligence, ergonomics, and human-computer interaction.

JUE QU received the B.S. and M.S. degrees from the Air Defense Anti-missile Academy, Air Force Engineering University, China. He is currently pursuing the Ph.D. degree with the School of Aeronautics, Northwestern Polytechnical University, China. He is also serves as an Associate Professor with Air Force Engineering University. He has published over 40 scientific articles in his research fields. His interests include optimization design, augmented reality, and ergonomics.