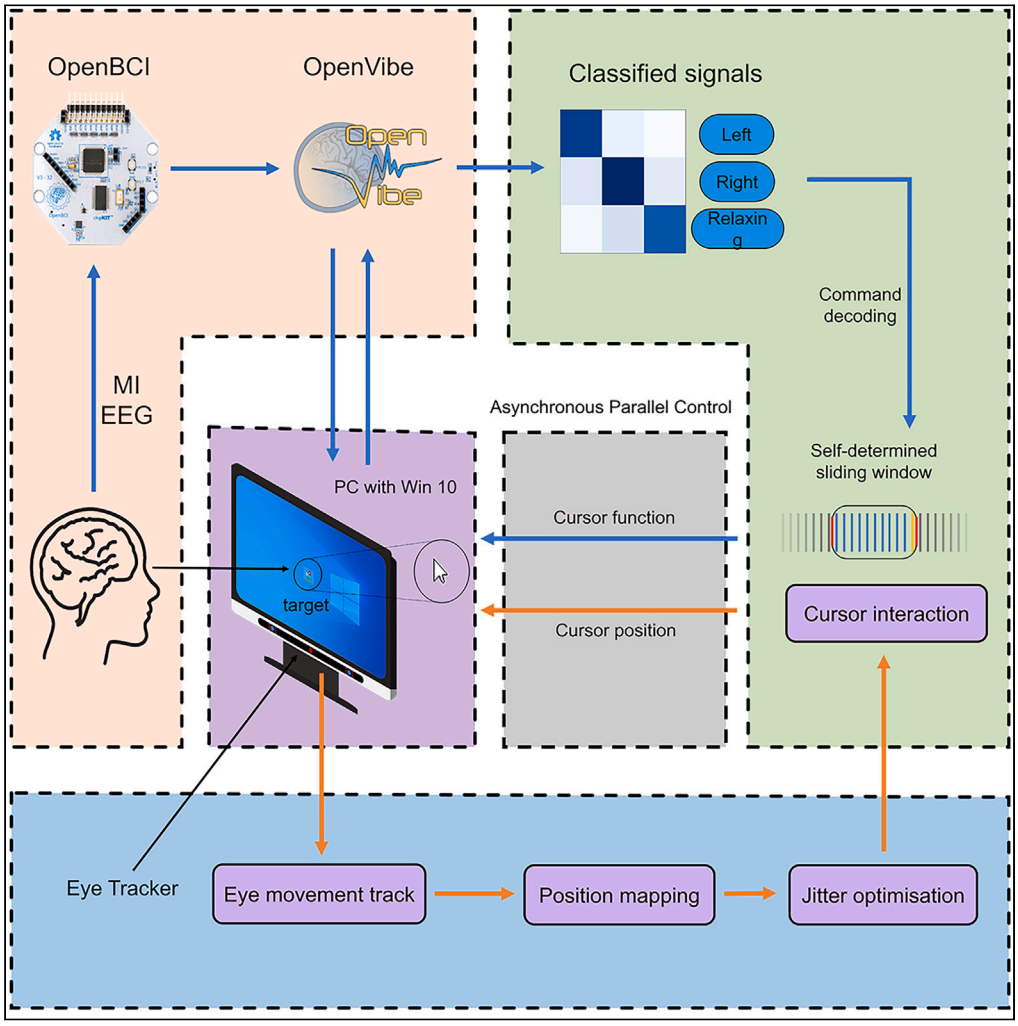
iScience

CellPress
OPEN ACCESS

Article

A high-performance general computer cursor control scheme based on a hybrid BCI combining motor imagery and eye-tracking



Jiakai Zhang, Yuqi Zhang, Xinlong Zhang, ..., Ju Wang, Shaojie Lu, Xiaoyan Shen

xiaoyansho@ntu.edu.cn

Highlights

A hybrid non-invasive BCI system for general computer cursor control

Using self-determined sliding window approach for command decoding

"Desire factor" optimization method improves the stability of the cursor

High performance in classification task and cursor typing task

Zhang et al., iScience 27, 110164
June 21, 2024
https://doi.org/10.1016/j.isci.2024.110164

iScience

CellPress
OPEN ACCESS

Article

A high-performance general computer cursor control scheme based on a hybrid BCI combining motor imagery and eye-tracking

Jiakai Zhang,[1] Yuqi Zhang,[1] Xinlong Zhang,[1] Boyang Xu,[1] Huanqing Zhao,[1] Tinghui Sun,[1] Ju Wang,[1] Shaojie Lu,[1] and Xiaoyan Shen[1,2,3,*]

SUMMARY

This study introduces a novel virtual cursor control system designed to empower individuals with neuromuscular disabilities in the digital world. By combining eye-tracking with motor imagery (MI) in a hybrid brain-computer interface (BCI), the system enhances cursor control accuracy and simplicity. Real-time classification accuracy reaches 87.92% (peak of 93.33%), with cursor stability in the gazing state at 96.1%. Integrated into common operating systems, it enables tasks like text entry, online chatting, email, web surfing, and picture dragging, with an average text input rate of 53.2 characters per minute (CPM). This technology facilitates fundamental computing tasks for patients, fostering their integration into the online community and paving the way for future developments in BCI systems.

INTRODUCTION

The development of the Internet is ever-changing. However, for patients with neuromuscular injuries, their inability to freely access information online means they struggle to keep up with the times. Physical disabilities not only impact their mental well-being but also introduce numerous inconveniences into their daily lives. Yet, the advent of brain-computer interface (BCI) technology offers a glimmer of hope for these individuals.[1,2] BCI, a fusion of neuroscience and engineering, facilitates direct communication between the human brain and computers, granting patients the ability to communicate with the outside world and control external devices.[3,4] Therefore, we hope that through this study, we can build a bridge for patients with neuromuscular injuries to rapidly integrate with the Internet society and provide patients with sufficient autonomy and freedom to contact the Internet, access the Internet, and utilize the Internet to have a positive psychological experience and enhance their quality of life.

Many BCI devices[5] as well as more efficient BCI training algorithms[6,7] have been used in real life. Currently, BCI can be used to implement functions such as drone control,[8] multi-directional wheelchair movement,[9] exoskeleton limb grasping,[10] and spellers.[11] However, the execution of these functions is frequently solitary, and the subjects cannot independently operate the computer control. Therefore, it is crucial to design a BCI system capable of implementing the mouse-control function. The preponderance of brain-controlled virtual cursor programs employs invasive brain-computer interfaces (iBCI) at present. Dekleva et al. introduced a novel decoding approach for cursor click based on identifying transient neural responses that emerge at the onset and offset of intended hand grasp. The transient-based approach offers high-functioning, generalized click control suitable for both point-and-click and click-and-drag applications.[12] Hosman et al. evaluated the decoding of intended cursor velocity from human motor cortical signals using a long short-term memory (LSTM) neural network (RNN). The experiment suggested that RNN decoding, when applied to human intracortical signals, could lead to significant improvements in continuous two-dimensional (2D) cursor control and can offer superior iBCI cursor control.[13] Although implantable brain-computer interfaces provide higher dimensional and temporal resolution of information, there are high surgical risks as well as high costs, and non-invasive solutions are more acceptable to patients compared to invasive solutions. Currently, there are non-invasive cursor control schemes realized with the help of Steady-State Visual Evoked Potential (SSVEP) and P300. Kapgate et al. explored the feasibility of using noninvasive hybrid SSVEP + P300 visual BCI (VBCI) for cursor control as a universal form of computer access. The proposed cursor control system has a graphical user interface (GUI) design that simultaneously evokes both SSVEP and P300 signals in the human cortex.[14] Hsu et al. utilized the high-frequency stimulation characteristics of SSVEP to study the feasibility of using iterative filtering-empirical mode decomposition (IF-EMD) to implement a BCI cursor system.[15] Wang et al. built a high-speed four-target BCI system for cursor control based on SSVEP, achieving the evaluation of vigilance levels in brain-computer interface tasks.[16] Li et al. introduce a hybrid BCI system leveraging P300 potentials and SSVEP for improved asynchronous control. This system, applied in real-time wheelchair control, shows enhanced accuracy and response speed through the innovative use of flickering buttons and intensified signals.[17] Additionally, the team provides a comprehensive review

[1]School of Information Science and Technology, Nantong University, Nantong 226019, China
[2]Nantong Research Institute for Advanced Communication Technologies, Nantong University, Nantong 226019, China
[3]Lead contact
*Correspondence: xiaoyansho@ntu.edu.cn
https://doi.org/10.1016/j.isci.2024.110164

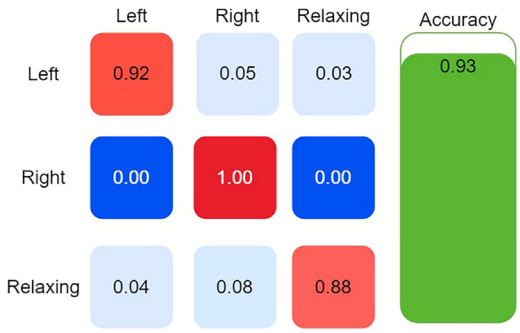


Figure 1. Confusion matrix for real-time three-class classification results for one of the subjects

of the latest developments in multimodal BCIs, highlighting several representative systems and their preliminary applications.[18] Although these methods fulfill the basic functions of a mouse, the subject's attention is fully devoted to the system's elicitation to the extent that he or she is unable to focus on what he or she really wants to view, and prolonged visual stimulation also causes eyestrain in the subject. The motor imagery (MI) electroencephalography (EEG) paradigm is a classic and extremely important paradigm in BCI research. By imagining the execution of an action in the subject's brain, the motor areas of the brain are activated, and non-invasive EEG equipment can detect this neural signal. This activity can be generated completely spontaneously and does not require external stimulation, making the use of MI EEG for cursor key function a good solution.

To compensate for the shortcomings of the motor imagery brain-computer interface paradigm of having fewer commands and at the same time achieve flexible control of the cursor, we introduce eye-tracker as an implementation solution. An eye-tracker is an instrument used to track the trajectory of eye movements and is widely used in gaming and cognitive domains. Currently in the field of BCI, Xu et al. achieved 3D control of multiple degrees of freedom (DOFs) robotic arms for multi-target reaching and grasping tasks with the help of a non-invasive EEG cap, an eye-tracker, and computer vision algorithms. In this case, eye-tracking was used to determine the target of attention, and then computer vision was used to find the target and grasp it.[19] Park et al. designed an AR-based home appliance control system for elderly end-users using SSVEP-based BCI and eye-tracking.[20] Li et al. proposed a practical hybrid BCI speller to improve the performance of Chinese character input by incorporating eye-tracking into the traditional P300 event-related potential spelling paradigm.[21] It is not difficult to find that the eye-tracker in the current brain-computer interface field only acts as an auxiliary cognitive judgment and selection tool. In addition, the need for an independent operating system and the problem of fewer functions are also reasons why the eye-tracker-based brain-computer interface system cannot be popularized at present. However, we believe that the problem of brain-controlled cursor orientation movement can be well solved if the eye movement trajectory can be mapped to the virtual cursor.

In this paper, a hybrid BCI system based on MI EEG signals and eye-tracking devices is proposed. The system takes the eye movement trajectory recorded by the eye-tracker as the movement strategy of the virtual cursor and maps "imagining right hand grasping" to the left click function of the cursor and "imagining left hand grasping" to the drag and drop function, thus realizing a hybrid BCI cursor system.

RESULTS

Results of the three-class classification experiment

Figure 1 depicts the confusion matrix for one of the subjects' real-time three-class classification results. The main diagonal of the confusion matrix represents the prediction of the three categories, whereas the columns and rows represent the actual and predicted categories, respectively, with darker colors indicating a more accurate prediction. In addition, we present in Table 1 the detailed three-category recognition results for the 16 subjects.

As can be seen in Table 1, the average recognition accuracy of the three classifications for the 16 subjects was (87.92 $\pm$ 2.67)%. For the three classes, the average recognition accuracies are as follows: (85.84 $\pm$ 5.62)% for "Imagining left hand grasping," (91.07 $\pm$ 3.95)% for "Imagining right hand grasping," and (86.84 $\pm$ 3.72)% for "Relaxing."

Gaze positioning accuracy results

In our study, we propose an optimization algorithm for the problem of cursor instability in the gaze state. To evaluate the efficacy of the algorithm, we designated a 64-pixel-diameter circle on the screen (1920 × 1080 pixels) as the gaze target. The subjects observed the target for 3 s while the system sampled at 90 Hz. In a comparison of each second as a unit, we present both the unoptimized and optimized outcomes. The experimental outcomes are illustrated in Figure 2.

As can be seen from the figure, the unoptimized gaze points are more discrete, and the optimized gaze points are more clustered. The stability is represented by the coverage of the sampled gaze points (blue dots) within the target area (red circles). The formula for calculation is as follows:

Table 1. Classification accuracies of MI for all 16 subjects in the online experiment

Subject	"Left" recognition accuracy (%)	"Right" recognition accuracy (%)	"Relaxing" recognition accuracy (%)	Total recognition accuracy (%)
1	92.13	100	87.87	93.33
2	94.05	86.6	91.92	90.86
3	86.39	87.61	85.94	86.65
4	90.2	88.1	79.85	86.05
5	91.7	90.11	85.22	89.01
6	87.01	87.53	90.05	88.20
7	78.93	92.51	84.96	85.47
8	90.63	91.16	87.4	89.73
9	84.21	94.7	88.27	89.06
10	76.09	89.11	83.82	83.01
11	86.01	97.88	78.99	87.63
12	85.49	88.33	90.56	88.13
13	86.81	95.11	89.68	90.53
14	85.11	90.77	90.52	88.80
15	84.98	89.41	85.62	86.67
16	73.76	88.23	88.84	83.61
AVG	85.84 ± 5.62	91.07 ± 3.95	86.84 ± 3.72	87.92 ± 2.67

$$Coverage\ Ratio = \left(\frac{Number\ of\ Points\ Inside\ the\ Circle}{Total\ Number\ of\ Points}\right) * 100\% \quad \text{(Equation 13)}$$

Since some of the sampling points are not entirely outside the target area (that is, they intersect with the red circles), these points also affect the stability. Therefore, we include these points in our calculations on a proportional basis. Quantifying the results gives us 72.9%, 75.5%, and 76.3% stability before optimization and 93.7%, 95.6%, and 96.1% stability after optimization.

The accuracy of the system was then evaluated during the transition between the eye movement state and the gaze state. Figure 3A depicts the scenario for testing. On a circle with 880 pixels, 16 smaller circles with 64 pixels are evenly distributed. The circle's center is the cursor's starting point, and the smaller circles in 16 directions are its destination locations. Subjects were required to begin from the center, return to the center of the circle after reaching the destination, and proceed clockwise to the next destination for 10 times. Figure 3B displays the results of the experiment.

Based on the experimental results, it is evident that this work can obtain a satisfactory cursor movement trajectory in 16 orientations and that the transformation of ocular states at the start and endpoints can also satisfy the requirements for daily computer operation.

Mouse control applications

The hybrid BCI operation scheme was applied to the Windows 10 operating system to verify the clicking and dragging functions.

Cursor typing application

Given the prevailing and prospective trend of utilizing BCI technology for text-based communication, it is noteworthy that several research teams have successfully implemented text input methods employing P300 or SSVEP BCI. In light of this, we conducted an experiment focused on cursor typing. The experiment may be conducted utilizing just the virtual keyboard provided by the Windows operating system and the left-click functionality employed in this investigation. The experimental findings are depicted in Figure 4.

Figure 4A shows the disordered typing with letter prompts. Figure 4B shows free typing without restriction. We can see that the subjects finished typing the sentence "Hello, nice to meet you!" in the Word document.

In the disordered typing experiment with alphabetic cues, the task consisted of 200 characters and 15% punctuation characters. The calculation of characters per minute (CPM) was conducted after completing an entire typing task. If an error was made during the task, the participant cannot continue until the error was corrected. However, the timer for the task continued to count. Thus, the process of error correction inevitably resulted in time loss, which impacted the overall CPM rate. The results showed an input accuracy of 93.97%, an average spelling speed of 53.2 CPM, and a word input speed of 24.03 words per minute. Table 2 shows the spelling speed comparison with different BCI spellers. It is evident that the velocity of typing utilizing the cursor in this task surpasses that of alternative methodologies.

The left-click operation is the most commonly utilized action in daily computer usage. Simultaneously, this operation holds paramount significance as it serves as the fundamental means of interacting with the computer. Hence, this study encompasses the many interactions

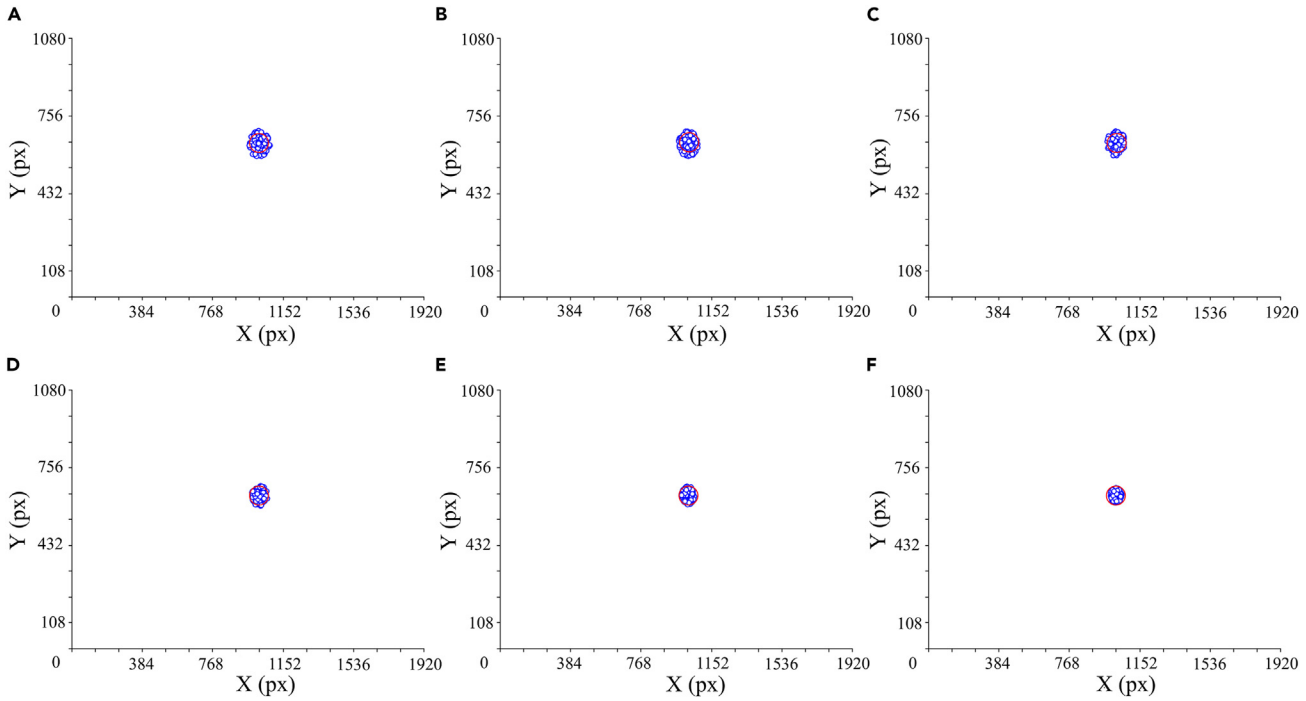


Figure 2. Comparison of jitter optimization results under 3-s gazing state

The red circles represent the gaze targets with a size of 64 pixels, whereas the blue dots indicate the sampled gaze points. The first, second, and third gaze sampling results are displayed from left to right, where (A–C) represent the unoptimized 3-s gaze point results, and (D–F) represent the optimized 3-s gaze point results.

facilitated by the left-click function of the mouse, including but not limited to, message transmission and reception, internet surfing, and email verification.

Drag-and-drop drawing application

In this experiment, we verified the drag-and-drop function of the cursor. We scattered several images on a canvas (Figure 5A), and subjects had to drag the images and combine them to form a picture containing the sun, a house, and a tree. Figure 5B is a list of experimental "artworks."

The average time taken by four participants to complete the painting is 45.45 s, with the fastest completion time being 39.23 s. The completion time for drawing is provided for reference only, yet the experimental results indicate that the work is also effective in executing the cursor's drag-and-drop functionality.

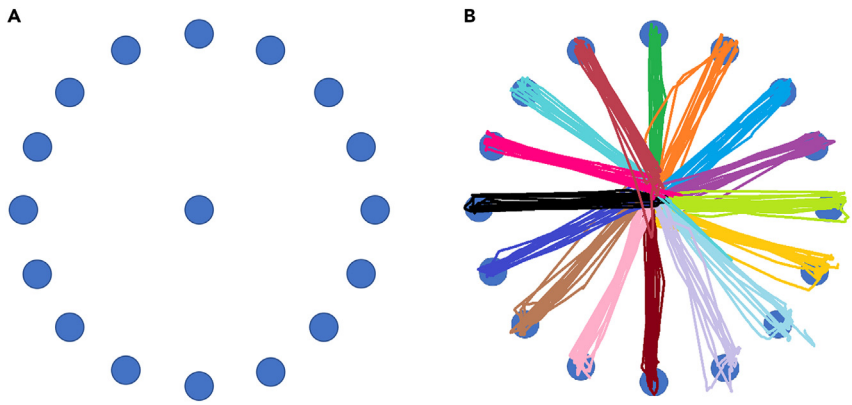


Figure 3. Test scheme and results for system accuracy during the transition between eye movement and gaze fixation states

(A) System test scheme with 16 small circles of 64 pixels evenly distributed on a circle of 880 pixels.

(B) System test results.

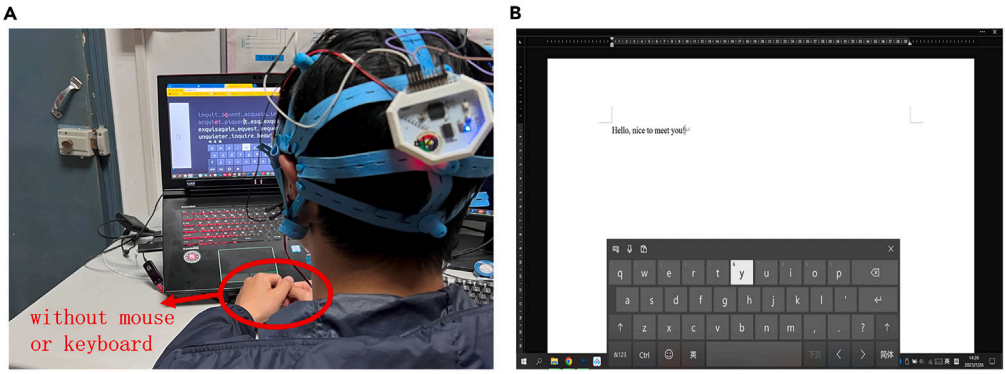


Figure 4. Cursor-typing experiment

(A) Subject performed a disordered typing experiment with alphabetic cues.

(B) Free-typing experiment.

DISCUSSION

In this study, the click-and-drag function of the virtual cursor was implemented using real-time MI EEG signal three-class classification and eye-tracking. The accuracy and functionality of the system were verified and analyzed by looking at the three-class classification accuracy, gaze positioning accuracy, cursor typing experiments, and drag-and-drop drawing experiments.

Three-class classification performance

In the three-class classification experiment, subjects 1, 2, and 13 had recognition accuracies of over 90%, and the overall average recognition accuracy for all 16 subjects was (87.92 $\pm$ 2.67)%. It is noteworthy that among the average recognition accuracy of all categories, the "imagining right hand grasping" category had a much higher recognition accuracy than the other two categories. One plausible explanation for this phenomenon is that the individuals involved in the study were right-handed, which therefore led to a heightened response while engaging in the motor imagery task of "imagining right hand grasping."

Gaze stability optimization

In the gaze positioning accuracy experiments, from the comparison of the vertical results, we can see that by adding the "desire factor," we can largely reduce the jittering problem of the cursor in the gaze state. From the horizontal results, it can be noticed that the stability is gradually increasing over the 3 s, whether optimization is performed or not. It is also worth noting that the gap in stability results between the first and second seconds is larger than the gap between the second and third seconds, a phenomenon that was observed both before and after the optimization. We see this as a process of gradual rise in attention. At the first second, the subjects' attention was just focused on the gaze target, and as time went on, the subjects' attention became more and more focused. This explains the gradual increase in stabilization over the 3 s. Once the attention is focused to a certain degree, the stabilization does not fluctuate as much, which explains why the stabilization gap is smaller in the last 2 s than in the first 2 s.

Typing experiment performance

In the disordered typing experiment with alphabetic cues, subjects completed the cued typing task well and completed free typing of sentences in the follow-up. During the cued typing phase, however, we observed that subjects made typographical errors. Because conventional blind typing relies more on muscle memory than on looking at the keyboard's keys, misspelled letters are common in daily life. However, in our cursor typing experiment, the virtual keyboard is presented on the screen. Therefore, subjects only need to press the corresponding keys on

Table 2. Spelling rate comparison with other types of BCI spellers

Team	Method	CPM (characters/min)
Lin et al.[22]	Tri-RSVP + P300	6
Hwang et al.[23]	SSVEPs	9.39
Nagel et al.[24]	Asynchronous VEPs	16.1
Nakanishi et al.[25]	SSVEPs	36.17
Mannan et al.[26]	SSVEP + Eye Tracking	25.79
Our	MI + Eye Tracking	53.2

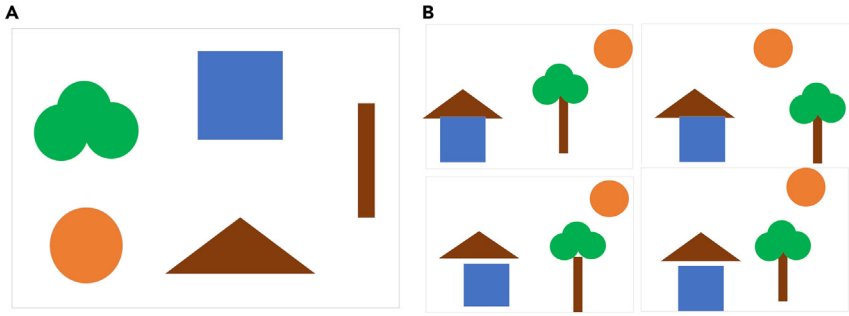


Figure 5. Drag-and-drop drawing experiment
(A) A canvas with scattered geometric shapes.
(B) Drag-and-drop drawing results.

the virtual keyboard according to the letter cues. This one-to-one correspondence task should be difficult to make mistakes on. After observing the experiment, we found that these errors were caused by the subjects' fast eye movements. In other words, after the subject issued the left click command, before clicking on the current letter, his eyes immediately moved to the next target letter, which led to mistouching. In the future, we will consider adding the function of key lock, i.e., locking the movement of the cursor during the clicking task, to achieve the purpose of precise interaction.

Limitations of the study

This research still has some opportunities for further refinement. The first is an occasional mismatch between cursor commands and interactive buttons due to rapid eye movements. The second is our current ability to recognize only a limited variety of EEG signal categories, which constrains our system's functional expansion. Importantly, our study primarily involved able-bodied subjects, indicating a gap in our understanding of the system's effectiveness for individuals with varying degrees of neuromuscular disabilities. This aspect highlights a critical area for future research, as thoroughly testing and adapting our system for users with different abilities is essential for its applicability. In our future research endeavors, we will continue to refine and enhance the applicability of our proposed method and system. Furthermore, we will delve into critical areas such as long-term usability in everyday environments, the effects of user fatigue, and the sustained reliability of the system over time. These focal points are pivotal for ensuring the practicality and efficacy of our system in real-world research settings.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - System architecture
  - Signal acquisition
  - EEG signal processing
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - BCI decoding performance
  - Gaze positioning accuracy performance

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.110164.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (Nos. 81371663 and 61534003), the "Six talents peaks" Project, China (No. SWYY-116), Nantong Natural Science Foundation (JC2023072).

AUTHOR CONTRIBUTIONS

Conceptualization: J.Z. and X.S.; methodology: J.Z. and B.X.; software: J.Z., B.X., H.Z., and J.W.; validation: J.Z., Y.Z., H.Z., T.S., and J.W.; visualization: J.Z.; data collection and curation: J.Z., Y.Z., X.Z., and B.X.; data analysis: J.Z., X.Z., and J.W.; investigation: T.S. and S.L.; recruitment: T.S. and S.L.; supervision: X.S.; funding acquisition: X.S.; writing—original draft: J.Z.; writing—review & editing: J.Z., Y.Z., and X.S.; project administration: X.S.

DECLARATION OF INTERESTS

The authors declare no competing interest.

Received: October 25, 2023
Revised: March 21, 2024
Accepted: May 29, 2024
Published: May 31, 2024

REFERENCES

1. Värbu, K., Muhammad, N., and Muhammad, Y. (2022). Past, present, and future of EEG-based BCI applications. Sensors *22*, 3331. https://doi.org/10.3390/s22093331.
2. Zhu, S., Yu, T., Xu, T., Chen, H., Dustdar, S., Gigan, S., Gunduz, D., Hossain, E., Jin, Y., Lin, F., et al. (2023). Intelligent computing: The latest advances, challenges, and future. Intell. Comput. *2*, 0006. https://doi.org/10.34133/icomputing.0006.
3. Mane, R., Chouhan, T., and Guan, C. (2020). BCI for stroke rehabilitation: motor and beyond. J. Neural. Eng. *17*, 041001. https://doi.org/10.1088/1741-2552/aba162.
4. Khan, M.A., Das, R., Iversen, H.K., and Puthusserypady, S. (2020). Review on motor imagery based BCI systems for upper limb post-stroke neurorehabilitation: From designing to application. Comput. Biol. Med. *123*, 103843. https://doi.org/10.1016/j.compbiomed.2020.103843.
5. Zhang, J., Li, J., Huang, Z., Huang, D., Yu, H., and Li, Z. (2023). Recent progress in wearable brain–computer interface (BCI) devices based on electroencephalogram (EEG) for medical applications: a review. Health Data Sci. *3*, 0096. https://doi.org/10.34133/hds.0096.
6. Pan, L., Wang, K., Xu, L., Sun, X., Yi, W., Xu, M., and Ming, D. (2023). Riemannian geometric and ensemble learning for decoding cross-session motor imagery electroencephalography signals. J. Neural. Eng. *20*, 066011. https://doi.org/10.1088/1741-2552/ad0a01.
7. Xie, Y., Wang, K., Meng, J., Yue, J., Meng, L., Yi, W., Jung, T.P., Xu, M., and Ming, D. (2023). Cross-dataset transfer learning for motor imagery signal classification via multi-task learning and pre-training. J. Neural. Eng. *20*, 056037. https://doi.org/10.1088/1741-2552/acfe9c.
8. Kim, S., Lee, S., Kang, H., Kim, S., and Ahn, M. (2021). P300 brain–computer interface-based drone control in virtual and augmented reality. Sensors *21*, 5765. https://doi.org/10.3390/s21175765.
9. Pawuś, D., and Paszkiel, S. (2022). BCI wheelchair control using expert system classifying EEG signals based on power spectrum estimation and nervous tics detection. Appl. Sci. *12*, 10385. https://doi.org/10.3390/app122010385.
10. Missiroli, F., Barsotti, M., Leonardis, D., Gabardi, M., Rosati, G., and Frisoli, A. (2019). Haptic stimulation for improving training of a motor imagery BCI developed for a hand-exoskeleton in rehabilitation. In 2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR) (IEEE), pp. 1127–1132. https://doi.org/10.1109/ICORR.2019.8779370.
11. Kundu, S., and Ari, S. (2022). Brain-computer interface speller system for alternative communication: a review. IRBM *43*, 317–324. https://doi.org/10.1016/j.irbm.2021.07.001.
12. Dekleva, B.M., Weiss, J.M., Boninger, M.L., and Collinger, J.L. (2021). Generalizable cursor click decoding using grasp-related neural transients. J. Neural. Eng. *18*, 0460e9. https://doi.org/10.1101/2020.09.03.20186973.
13. Hosman, T., Vilela, M., Milstein, D., Kelemen, J.N., Brandman, D.M., Hochberg, L.R., and Simeral, J.D. (2019). BCI decoder performance comparison of an LSTM recurrent neural network and a Kalman filter in retrospective simulation. In 2019 9th International IEEE/EMBS conference on neural engineering (NER) (IEEE), pp. 1066–1071. https://doi.org/10.1109/NER.2019.8717140.
14. Kapgate, D. (2022). Effective 2-D cursor control system using hybrid SSVEP + P300 visual brain computer interface. Med. Biol. Eng. Comput. *60*, 3243–3254. https://doi.org/10.1007/s11517-022-02675-0.
15. Hsu, C.C., Yeh, C.L., Lee, W.K., Hsu, H.T., Shyu, K.K., Li, L.P.H., Wu, T.,Y., and Lee, P.L. (2020). Extraction of high-frequency SSVEP for BCI control using iterative filtering based empirical mode decomposition. Biomed. Signal Process Control *61*, 102022. https://doi.org/10.1016/j.bspc.2020.102022.
16. Wang, K., Qiu, S., Wei, W., Zhang, C., He, H., Xu, M., and Ming, D. (2021). Vigilance estimating in SSVEP-based BCI using multimodal signals. In 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (IEEE), pp. 5974–5978. https://doi.org/10.1109/embc46164.2021.9629736.
17. Li, Y., Pan, J., Wang, F., and Yu, Z. (2013). A hybrid BCI system combining P300 and SSVEP and its application to wheelchair control. IEEE Trans. Biomed. Eng. *60*, 3156–3166. https://doi.org/10.1109/TBME.2013.2270283.
18. Li, Y., Pan, J., Long, J., Yu, T., Wang, F., Yu, Z., and Wu, W. (2015). Multimodal BCIs: target detection, multidimensional control, and awareness evaluation in patients with disorder of consciousness. Proc. IEEE *104*, 332–352. https://doi.org/10.1109/JPROC.2015.2469106.
19. Xu, B., Li, W., Liu, D., Zhang, K., Miao, M., Xu, G., and Song, A. (2022). Continuous hybrid BCI control for robotic arm using noninvasive electroencephalogram, computer vision, and eye tracking. Mathematics *10*, 618–623. https://doi.org/10.3390/math10040618.
20. Park, S., Ha, J., Park, J., Lee, K., and Im, C.H. (2022). Brain-controlled, AR-based Home automation system using SSVEP-based brain-computer interface and EOG-based eye tracker: A feasibility study for the elderly end User. IEEE Trans. Neural Syst. Rehabil. Eng. *31*, 544–553. https://doi.org/10.1109/TNSRE.2022.3228124.
21. Li, H., Fang, X., Ye, Z., Chu, X., Lu, G., and Yu, Y. (2022). A practical hybrid BCI speller for Chinese Character Input: Integrating an Eye Tracker into a P300-Based Spelling approach. In Proceedings of the 5th International Conference on Computer Science and Software Engineering (Association for Computing Machinery), pp. 237–242. https://doi.org/10.1145/3569966.3570046.
22. Lin, Z., Zhang, C., Zeng, Y., Tong, L., and Yan, B. (2018). A novel P300 BCI speller based on the Triple RSVP paradigm. Sci. Rep. *8*, 3350. https://doi.org/10.1038/s41598-018-21717-y.
23. Hwang, H.J., Lim, J.H., Jung, Y.J., Choi, H., Lee, S.W., and Im, C.H. (2012). Development of an SSVEP-based BCI spelling system adopting a QWERTY-style LED keyboard. J. Neurosci. Methods *208*, 59–65. https://doi.org/10.1016/j.jneumeth.2012.04.011.
24. Nagel, S., and Spüler, M. (2019). Asynchronous non-invasive high-speed BCI speller with robust non-control state detection. Sci. Rep. *9*, 8269. https://doi.org/10.1038/s41598-019-44645-x.
25. Nakanishi, M., Wang, Y., Chen, X., Wang, Y.T., Gao, X., and Jung, T.P. (2018). Enhancing detection of SSVEPs for a high-speed brain speller using task-related component analysis. IEEE Trans. Biomed. Eng. *65*, 104–112. https://doi.org/10.1109/TBME.2017.2694818.
26. Mannan, M.M.N., Kamran, M.A., Kang, S., Choi, H.S., and Jeong, M.Y. (2020). A hybrid speller design using eye tracking and SSVEP brain–computer interface. Sensors *20*, 891. https://doi.org/10.3390/s20030891.
27. Lotze, M., and Halsband, U. (2006). Motor imagery. J. Physiol. Paris *99*, 386–395. https://doi.org/10.1016/j.jphysparis.2006.03.012.

28. Blanchard, G., and Blankertz, B. (2004). BCI competition 2003-data set IIa: spatial patterns of self-controlled brain rhythm modulations. IEEE Trans. Biomed. Eng. *51*, 1062–1066. https://doi.org/10.1109/TBME.2004.826691.

29. Cheng, M., Jia, W., Gao, X., Gao, S., and Yang, F. (2004). Mu rhythm-based cursor control: an offline analysis. Clin. Neurophysiol. *115*, 745–751. https://doi.org/10.1016/j.clinph.2003.11.038.

30. Pfurtscheller, G. (2001). Functional brain imaging based on ERD/ERS. Vis. Res. *41*, 1257–1260. https://doi.org/10.1016/S0042-6989(00)00235-2.

31. Khazi, M., Kumar, A., and Vidya, M.J. (2012). Analysis of EEG using 10: 20 electrode system. Int. J. Innov. Res. Sci. Eng. Technol. *1*, 185–191.

32. Angrisani, L., Arpaia, P., De Benedetto, E., Duraccio, L., Regio, F.L., and Tedesco, A. (2023). Expanding the Frontiers of Wearable Brain-Computer Interfaces Combining Augmented Reality and Visually Evoked Potentials. In 2023 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE) (IEEE), pp. 58–62. https://doi.org/10.1109/MetroXRAINE58569.2023.10405702.

33. Selim, S., Tantawi, M.M., Shedeed, H.A., and Badr, A. (2018). A csp\am-ba-svm approach for motor imagery bci system. IEEE Access *6*, 49192–49208. https://doi.org/10.1109/ACCESS.2018.2868178.

34. Park, Y., and Chung, W. (2019). Selective feature generation method based on time domain parameters and correlation coefficients for Filter-Bank-CSP BCI systems. Sensors *19*, 3769. https://doi.org/10.3390/s19173769.

35. Zhang, J., Wang, X., Xu, B., Wu, Y., Lou, X., and Shen, X. (2023). An overview of methods of left and right foot motor imagery based on Tikhonov regularisation common spatial pattern. Med. Biol. Eng. Comput. *61*, 1047–1056. https://doi.org/10.1007/s11517-023-02780-8.

36. Zhang, J., Xu, B., Lou, X., Wu, Y., and Shen, X. (2023). MI-based BCI with accurate real-time three-class classification processing and light control application. Proc. Inst. Mech. Eng. H *237*, 1017–1028. https://doi.org/10.1177/09544119231187287.

37. Arpaia, P., Esposito, A., Natalizio, A., and Parvis, M. (2022). How to successfully classify EEG in motor imagery BCI: A metrological analysis of the state of the art. J. Neural. Eng. *19*, 031002. https://doi.org/10.1088/1741-2552/ac74e0.

38. Pisner, D.A., and Schnyer, D.M. (2020). Support vector machine. In Machine learning: Methods and applications to brain disorders, A. Mechelli and S. Vieira, eds. (Academic Press), pp. 101–121. https://doi.org/10.1016/B978-0-12-815739-8.00006-7.

39. Olsen, A. The Tobii I-VT Fixation Filter: Algorithm Description. https://scholar.google.se/citations?view_op=view_citation&hl=en&user=cAHHyyMAAAAJ&citation_for_view=cAHHyyMAAAAJ:u5HHmVD_uO8Chttp://www.vinis.co.kr/ivt_filter.pdf.

40. Mughrabi, M.H., Mutasim, A.K., Stuerzlinger, W., and Batmaz, A.U. (2022). My Eyes Hurt: Effects of Jitter in 3D Gaze Tracking. In 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) (IEEE), pp. 310–315. https://doi.org/10.1109/VRW55335.2022.00070.

41. Yershov, R., Voytenko, V., and Bychko, V. (2019). Software-based contact debouncing algorithm with programmable auto-repeat profile feature. In 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T) (IEEE), pp. 813–818. https://doi.org/10.1109/PICST47496.2019.9061500.

42. Gaur, P., Gupta, H., Chowdhury, A., McCreadie, K., Pachori, R.B., and Wang, H. (2021). A sliding window common spatial pattern for enhancing motor imagery classification in EEG-BCI. IEEE Trans. Instrum. Meas. *70*, 1–9. https://doi.org/10.1109/TIM.2021.3051996.

43. Zhao, Q., Zhang, L., and Cichocki, A. (2009). EEG-based asynchronous BCI control of a car in 3D virtual reality environments. Chin. Sci. Bull. *54*, 78–87. https://doi.org/10.1007/s11434-008-0547-3.

44. Ma, C., Li, W., Cao, J., Du, J., Li, Q., and Gravina, R. (2020). Adaptive sliding window based activity recognition for assisted livings. Inf. Fusion *53*, 55–65. https://doi.org/10.1016/j.inffus.2019.06.013.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
rowheadDeposited data		
Dataset	Authors	https://github.com/eytydgshib/MI-BCI-real.git
rowheadSoftware and algorithms		
MATLAB	MathWorks	https://www.mathworks.com/
Visual Studio 2013	Microsoft	https://visualstudio.microsoft.com/zh-hans/
Python	Python Software Foundation	https://www.python.org/
Tobii SDK	Tobii AB	https://developer.tobiipro.com/
OpenViBE	Inria Rennes	http://openvibe.inria.fr/
rowheadOther		
OpenBCI Cyton	OpenBCI	https://openbci.com/
Tobii 4C	Tobii AB	https://gaming.tobii.com/products

RESOURCE AVAILABILITY

Lead contact

Xiaoyan Shen (xiaoyansho@ntu.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Data is available from the lead contact upon request.
- The code is publicly available at https://github.com/eytydgshib/MI-BCI-real.git.
- Any additional information required is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

A total of sixteen participants were recruited for the study (10 males and 6 females, average age: 24.6 $\pm$ 1.49 years). All subjects were right-handed and had normal or corrected vision. In addition, all subjects had no history of psychological or neurological disorders. All subjects gave informed consent to the experimental study and signed an informed consent form. The experimental procedures were approved by the Ethics Committee of Nantong University (NTU Ethics Approval (2022) No. 1). All participant data and privacy will be strictly protected throughout the study. A description of the participants characteristics is provided in the Table S1.

METHOD DETAILS

System architecture

The system includes a computer, an EEG collection unit, an eye tracking unit, and a signal command processing unit, where the signal command processing unit is implemented on the same computer as the experimental presentation of the cursor control. We have divided the cursor control functionality into two distinct operations: key functions and movement functions.

For the cursor key function, subjects imagined performing left- and right-hand grasping and relaxing movements, and the MI EEG signals generated by this process were captured and processed by a computer to obtain the final three-class classification results. The classification results are then corresponded to the cursor key commands. Given the prevalence of mouse key utilization and the dominant hand preference, we establish a correlation whereby the "Imagining right hand grasping" is associated with the left click function, while the "Imagining left hand grasping" is linked to the drag and drop function. To mitigate control mistakes resulting from sporadic erroneous signals and enhance the precision and comfort of system control, we proposed a self-determined sliding window approach for command decoding of EEG signals.

To facilitate the cursor movement function, we activated an eye tracker to capture the gaze point coordinates of participants in real time, thereby acquiring the cursor's position information. Subsequently, the gaze point is converted into pixel coordinates and stored in the buffer. The process of mapping the cursor position can be achieved by extracting the coordinates from the buffer and afterward substituting them with the coordinates of the original cursor location. Subsequently, jitter optimization of the cursor in the gaze state enables cursor following.

Participants will be asked to complete a set of target transitions/gaze tasks in 16 directions. During this period, the system will record the cursor tracking trajectory to help participants become familiar with controlling the cursor using eye movements in the initial phase.

Finally, the cursor clicking functions controlled by MI EEG and the cursor movement functions controlled by eye tracking operate on the computer cursor in an asynchronous parallel manner, achieving complete cursor functionality. Due to the asynchronous parallel control, the clicking and movement functions do not conflict with each other, effectively replicating the operational logic of a real mouse. In practical applications, participants move the cursor via eye movements and execute clicking commands at the desired target location through MI, completing a cursor operation just as they would with a normal mouse (with the timing of execution entirely determined by the participant).

Signal acquisition

EEG signal acquisition module

The EEG signal acquisition device utilized in this study was OpenBCI (OpenBCI, New York, NY, USA), which operated at a sample frequency of 250 Hz. Upon receiving instructions to perform MI tasks, participants exhibited activation in the pertinent regions of the cerebral cortex, which ultimately led to the production of EEG signals displaying regular variations. The physiological rationale for utilizing MI EEG signals as input to the BCI system lies in the notable stability and variability exhibited by these signals during specified temporal intervals.[27] When unilateral limb movements were imagined, the μ (8–12 Hz) and β (13–30 Hz) rhythms in the contralateral primary sensorimotor cortical areas of the brain exhibited a considerable decrease in energy.[28,29] This is known as event-related desynchronization (ERD). In contrast, the energy of rhythms in the ipsilateral primary sensorimotor cortical regions significantly increased. This result is known as event-related synchronization (ERS)[30] Due to the concentration of EEG signals related to the mental imagery of left-right movements in the C3 and C4 regions, the study employed a simplified approach by utilizing only three dry electrodes - C3, Cz, and C4 - as measurement electrodes to alleviate the challenges associated with processing multichannel data and enhance the overall comfort experienced by the participants. The electrodes were positioned in accordance with the internationally recognized 10–20 standard for electrode placement, [31] as depicted in Figure S1. Additionally, two ear clip electrodes were utilized as reference electrodes.

OpenViBE (http://openvibe.inria.fr) is a free and open-source software for the recognition and processing of EEG signals in real time. The Acquisition Server of OpenBCI and OpenViBE were interconnected to acquire real-time, high-resolution brain activity information. Subsequently, the MI BCI stimulator was accessed to provide stimulus cues to the participants. The stimulus cues were classified into three categories: left, right, and up arrows. The left and right arrows were used to symbolize the motor imagery of left- and right-hand grasping, respectively, while the up arrow denoted relaxing. All three arrows appeared randomly and at the same interval. The participants engaged in the respective motor imagery task in response to the provided stimulus cues, while the computer captured and documented the status of their brain activity. The experimental paradigm is displayed in Figure S2. Prior to the experiment, all participants were provided with ample time to become acquainted with the EEG cap and the experimental procedure. This process required 10–15 min before the formal experiment. During the training session, every participant was instructed to conduct a total of 10 sets of tests, and each experiment consisted of three left-right MI tasks and three relaxing tasks. Before each experiment begins, participants have 1 s to prepare. Then, a green cross appears on the screen for 2 s to help the participants focus their attention. Subsequently, the system prompts the participants and records the EEG signals at that moment, continuing for 5 s.

Mapping of gaze point to cursor position

This study employed eye tracking technology to ascertain the position of the cursor during the online task. An off-screen eye tracker, Tobii 4C (Tobii Tech, Stockholm, Sweden), was used to detect the subject's gaze position and map its coordinates to a computer cursor to control the cursor on the monitor in front of the subject (a Twisted Nematic (TN) display with a resolution of 1920 * 1080 pixels and a ratio of 16:9 was chosen for this work).

The gaze point acquired using the eye-tracker was not a precise point, but rather fell inside a circular area with a diameter of 96 pixels. Hence, to perform cursor mapping, we choose to utilize the coordinates corresponding to the center of the circle and afterward normalized such coordinates. The coordinate system used has its origin at the bottom left corner of the screen, denoted as (0, 0), and extends to the top right corner, represented as (1, 1). The outcomes of the processing are visually depicted in Figure S3.

Since the coordinates of the cursor are represented in the form of pixel coordinates. Hence, it is imperative to transform the gaze coordinates into pixel coordinates in a proportionate manner. To perform the mapping process, it is necessary to acquire the parameters of the experimental screen. This involves determining the proportion of (x_0, y_0) within the X axis and Y axis value ranges. Subsequently, these proportions are multiplied by the pixel resolution of the display to derive the corresponding pixel coordinates. As an illustration, where the coordinates of the gaze point are (0.662, 0.437), the appropriate pixel coordinates may be calculated as follows:

$$\left(\frac{0.662}{1} * 1920px, \frac{0.437}{1} * 1080px\right) = (1271px, 472px) \quad \text{(Equation 1)}$$

Ultimately, the pixel coordinate is employed as the resultant value to supplant the existing data pertaining to the cursor position.

The process of mapping from gaze point to cursor position is shown in Figure S4. The initial step involves calibrating the eye movement data of the participant. Subsequently, the real-time original gaze point coordinates of the participant are derived by fitting the eye movement data with the calibration information. Then, the acquired original gaze point coordinates are normalized. Finally, the normalized gaze point

coordinates are converted to pixel coordinates and recorded into the buffer. Checking for existing viewpoint coordinates in the buffer before cursor mapping is the initial step; if not, return to the previous step to obtain the viewpoints and rewrite them. The data will be replaced with the coordinates of the current cursor if there are already coordinates in the buffer, in order to realize positional mapping and empty the cache in order to free up memory.

EEG signal processing

EEG Preprocessing

Wavelet Transform filters are applied to the signals at the electrodes, aimed at enhancing the signals of relevant brain regions through spatial differencing. Following this, an 8th-order Butterworth band-pass filter is used to filter the EEG signals, with the frequency ranges set as 4–8 Hz, 8–12 Hz, 12–16 Hz, 16–20 Hz, 20–24 Hz, 24–28 Hz, 28–32 Hz, 32–36 Hz, and 36–40 Hz. This is done to precisely analyze the signal quality of each sub-band and generate specific filter characteristics.[32] By meticulously evaluating and processing the signals of each frequency band, we ensure that only high-quality data are used for the final classification analysis.

MI-EEG signal feature extraction

During the process of human motor imagery, the brain generates two distinct types of rhythmic signals that exhibit noticeable variations: the mu-rhythm signal, which oscillates at a frequency range of 8–12 Hz, and the beta-rhythm signal, which oscillates at a frequency range of 13–30 Hz[28] The presence of ERS and ERD phenomena can lead to alterations in the power spectra of mu-rhythm and beta-rhythm in EEG data.[29,30] Hence, spatial filters can be employed for the purpose of feature extraction in relation to this energy shift. The common spatial pattern (CSP) method is well recognized as the predominant spatial filtering technique in use.[33,34] The algorithm leverages the spatial distribution of features to project the EEG signal onto a subspace. By diagonalizing the matrix, the algorithm identifies a set of spatial filters that optimize the projection process. These filters aim to maximize the variance difference between two types of signals, thereby generating feature vectors that exhibit the desired level of differentiation. The CSP algorithm is expressed as follows:

$$J(\omega) = \frac{\omega^T X_1^T X_1 \omega}{\omega^T X_2^T X_2 \omega} = \frac{\omega^T C_1 \omega}{\omega^T C_2 \omega}, \quad \text{(Equation 2)}$$

where ω is the spatial filter trained from the extracted feature vectors, T represents the transpose, X_i represents the data matrix of the motion imagery classes, and C_i represents the covariance matrix of classes 1 and 2. Although the CSP algorithm can gain algorithmic advantages in time and space, there is an overfitting problem when the data sample is too small.

In this paper, the Tikhonov regularized common spatial pattern (TRCSP) method is employed as a solution to address the issue of overfitting.[35,36] The regularization of the CSP algorithm, as shown in Equation 1, is achieved by imposing a prior on the spatial filter. These prior serves to penalize solutions that do not adhere to the prior information. The penalty function $P(\omega)$ in Equation 2 is the metric employed to quantify the extent to which the spatial filter ω adheres to the specified previous condition. As the degree of requirements that ω fulfills increases, the magnitude of $P(\omega)$ decreases. The TRCSP method is a fusion of the CSP and Tikhonov regularization (TR) techniques. The TR model is a well-established regularization technique that was initially proposed for regression issues to address the issue of including penalty terms through the use of weights. The penalty term $P(\omega)$ for the TRCSP method is defined as follows:

$$J(\omega) = \frac{\omega^T C_1 \omega}{\omega^T C_2 \omega + \alpha P(\omega)}, \quad \text{(Equation 3)}$$

$$P(\omega) = \|\omega\|^2 = \omega^T \omega = \omega^T I \omega, \quad \text{(Equation 4)}$$

where α in Equation 2 is the regularization parameter which can be defined by us (the higher α, the more satisfied the prior), and I in Equation 3 represents the identity matrix. The optimization process of changing a spatial filter can yield good results even when there is minimal training data or noise present. After several iterations, the parameter "α" was ultimately assigned a value of 3.

One notable feature of the TRCSP method is that the newly formulated objective function remains a generalized eigenvalue problem, which can be solved by solving the eigenvalue problem. Nevertheless, due to the lack of symmetric features in the numerator and denominator of the objective function after including the regular component, it is necessary to get two eigenvalue solutions to derive the ultimate spatial filter. Specifically, we need to solve the following two objective functions:

$$J_1(\omega) = \frac{\omega^T C_1 \omega}{\omega^T C_2 \omega + \alpha \omega^T I \omega}, \quad \text{(Equation 5)}$$

$$J_2(\omega) = \frac{\omega^T C_2 \omega}{\omega^T C_1 \omega + \alpha \omega^T I \omega}, \quad \text{(Equation 6)}$$

The maximization of $J_1(\omega)$ corresponds to the maximization of variance in the first class of tasks and the minimization of variance in the second class of tasks. Conversely, the maximization of $J_2(\omega)$ corresponds to the maximization of variance in the second class of tasks and

the minimization of variance in the first class of tasks. The matrices of eigenvectors produced by the solution of the objective function are as follows:

$$M_1 = (C_2+\alpha I)^{-1}C_1, \quad \text{(Equation 7)}$$

$$M_2 = (C_1+\alpha I)^{-1}C_2, \quad \text{(Equation 8)}$$

The ideal spatial filter is formed by selecting the feature vectors that correspond to the greatest eigenvalues. The ensuing feature extraction procedure is consistent with the classic CSP approach.

Pattern recognition

Pattern recognition refers to the cognitive process of extracting distinctive characteristics from a given dataset through the acquisition of input signals. These acquired features are subsequently utilized to categorize, recognize, segment, or predict novel data points. The outcome of pattern recognition serves as the primary and immediate indication of the subject's cognitive state in BCI systems. Hence, the utilization of an efficient and precise pattern recognition methodology is of utmost importance.

The Support Vector Machine (SVM) is a widely employed supervised learning method that is frequently utilized for the resolution of binary and multiple classification issues.[37,38] By mapping the samples into a high-dimensional feature space and utilizing the support vectors in the feature space, the SVM finds the hyperplane. The objective of optimization is to identify an ideal hyperplane that simultaneously maximizes the sample margin and minimizes the sample error. SVM exhibit quicker training times in comparison to neural network methods. This advantage stems from the ability of SVM to turn the optimization issue into a convex optimization problem, allowing for the attainment of the global optimal solution using an efficient mathematical technique.

In our study, three types of EEG signals need to be categorized. The problem can be decomposed into three binary classification subproblems. In each binary classification problem, one category is treated as positive and the others as negative, thus constituting a "one-to-rest" support vector machine, as shown in Figure S5. Three sets of OVR–SVM classifiers were "left vs. right and relaxing," "right vs. left and relaxing," and "relaxing vs. left and right." The ratio of training and testing sets was 7:3. Only the results of the positive set were obtained for each set of classifiers $f_i(x), x = 1,2,3$, and the result files were then combined into a single triple classification output. In real-time classification, the largest $f_i(x)$ was used as the output.

Functional optimization

To attain system stability and accuracy during interactions, while mitigating mistakes resulting from sporadic erroneous signals, we proposed algorithmic optimization in both cursor jitter and command discrimination.

Cursor jitter optimization method

Cursor jitter is the erratic movement of the cursor on the display. Traditionally, the manipulation of cursors has relied on the utilization of external computer input devices, such as mice and touchpads. These external input devices utilize mechanical or optical principles to attain accurate and consistent control over the cursor. In contrast to conventional control techniques, the accuracy of gazing localization is influenced by both its inherent precision and other physiological conditions like visual fatigue, resulting in the occurrence of inevitable jitter. Optimizing the jitter of the cursor has the potential to boost operational precision, improve the user experience, and mitigate eye strain. Currently, several experts and researchers have put forward various optimization strategies to address the issue of jitter in the eye movement state.[39,40] However, it is worth noting that the interaction process predominantly takes place during the gaze state.

In this research, we present a novel approach termed the "desire factor" optimization method to address the issue of jitter in the gaze state. A "desire factor" γ is introduced into the cursor coordinate update expression. The specific value of γ is determined by evaluating the disparity between the current coordinates and the coordinates at the previous moment. Subsequently, γ is multiplied by the coordinate difference and added to the previous moment's coordinate value to obtain the updated cursor position. The expression for updating the coordinates is presented in Equations 8 and 9.

$$X_{new} = X_{pre} + \gamma_X(X_{curr} - X_{pre}), \quad \text{(Equation 9)}$$

$$Y_{new} = Y_{pre} + \gamma_Y(Y_{curr} - Y_{pre}), \quad \text{(Equation 10)}$$

where the variables X_{new} and Y_{new} represent the new cursor coordinates, whereas X_{pre} and Y_{pre} represent the coordinates at the previous instant. The variables X_{curr} and Y_{curr} represent the current coordinates. Additionally, γ_X and γ_Y represent the desire factor in the X-direction and Y-direction, respectively. The value of the "desire factor" γ is determined by the following functional relationship:

$$\gamma_X = \frac{1}{1+e^{\left(\frac{48 - A_X}{20}\right)}}, A_X = |X_{curr} - X_{pre}|, \quad \text{(Equation 11)}$$

$$\gamma_Y = \frac{1}{1+e^{\left(\frac{48 - A_Y}{20}\right)}}, A_Y = \left|Y_{curr} - Y_{pre}\right|, \quad \text{(Equation 12)}$$

where the variable A denotes the absolute value of the difference between the current coordinate and the preceding coordinate. The formula has the capacity to indicate the magnitude of the "desire" of the cursor to undergo movement. When the magnitude of variable A increases, it indicates a higher level of intention to move the cursor, resulting in a value of γ that approaches 1. Conversely, when the magnitude of A decreases, it signifies a lower level of intention to move the cursor, leading to a value of γ that is near to 0.

In this study, the proposed jitter optimization algorithm based on the "desire factor" reduces the jitter of the cursor during gaze and greatly improves stability during interaction.

Cursor command decoding for self-determined sliding window

In our work, we translated MI EEG signals to cursor key instructions, mapping "imagining right hand grasping" to the left click function of the cursor and "imagining left hand grasping" to the drag and drop function. However, if the EEG signals sampled instantaneously are only outputted in an "if-else" manner, not only would the entire system run too quickly, but we would also run into the issue of producing incorrect commands when there isn't enough time to rectify or reject the error signals. This increases the mistake rate and leads to a rise in the system's error rate, which has a significant impact on the operational experience. One potential solution to address the issue of the system operating at an excessively high frequency is to introduce a time delay[41] However, it is important to note that the sampling frequency of the EEG cap remains unchanged. Consequently, the EEG cap continues to sample data during the system's delay period, resulting in the subsequent discarding of the sampled data due to the system's delay. Therefore, the delay strategy may significantly diminish the retention of valuable signals. Moreover, the utilization of sliding windows is a prevalent method for decoding. The conventional sliding window discrimination method employs a window with a fixed extent to separate signals. However, in our experiment, the commands provided by the participants were created autonomously and exhibited random characteristics.[42,43] As a result, the use of a fixed width sliding window was unable to promptly differentiate between these directives. Even though adaptive sliding-window decoding algorithms have been presented, [44] the modification of the window size is postponed until after the transmission of the signal stream, resulting in significant delays. Considering the uncertainty of the command issued in this experiment, the strategy of window resizing cannot match the actual control. The adaptive sliding window decoding method is therefore inappropriate for this scenario.

In order to ensure more comfortable and accurate control of the system and to reduce the sporadic errors, we propose a self-determined sliding window approach to decoding the commands. The term "self-determined" in the context is intended to indicate that the sliding window's durations and overlaps for command decoding are adapted based on the individual's MI EEG signal characteristics rather than being fixed *a priori*. This approach is somewhat akin to self-paced or asynchronous MI window selection, where the system dynamically adjusts to the user's natural variability in executing MI tasks, rather than enforcing a strict, uniform timing across all users. The features "imagining left/right hand grasping" corresponded to the starting pointer, denoted as "left-b/right-b" of the sliding window. Similarly, the feature "relaxing" was used to describe the ending pointer, denoted as "relax-e" of the sliding window. Given the prevalence of mouse key utilization and the dominant hand preference, we establish a correlation whereby the "Imagining right hand grasping" is associated with the left click function, while the "Imagining left hand grasping" is linked to the drag and drop function, and "relaxing" signal take no action. As shown in Figure S6, when "imagining right hand grasping" is detected, the signal flow is marked with "right-b" as the beginning of the sliding window; when "relaxing" is received, the signal flow is marked with "relax-e" as the ending of the sliding window. This method returns a sliding window with "right-b" as the starting pointer and "relax-e" as the ending pointer; the output of this sliding window corresponds to the left-click method. Similarly, starting with "imagining left-handed grasping" and ending with "relaxing" will result in a sliding window starting at "left-b" and ending at "relax-e", whose output corresponds to a drag and drop function, where dragging starts at "left-b" and ends at "relax-e".

QUANTIFICATION AND STATISTICAL ANALYSIS

BCI decoding performance

EEG data from the subjects were exported and analyzed using OpenViBE 2.2.0 and Excel 2021 to determine the classification accuracy for each category of motor imagery (Table 1).

Gaze positioning accuracy performance

Eye-tracking sampling points within a 3-s interval were obtained using Python 3.8.10, and the stability was calculated. The stability was represented by the coverage ratio of the sampling points within the target area, using the following formula:

$$Coverage\ Ratio = \left(\frac{Number\ of\ Points\ Inside\ the\ Circle}{Total\ Number\ of\ Points}\right) * 100\%$$

Since some of the sampling points are not entirely outside the target area (that is, they intersect with the red circles), these points also affect the stability. Therefore, we include these points in our calculations on a proportional basis (Figure 2).