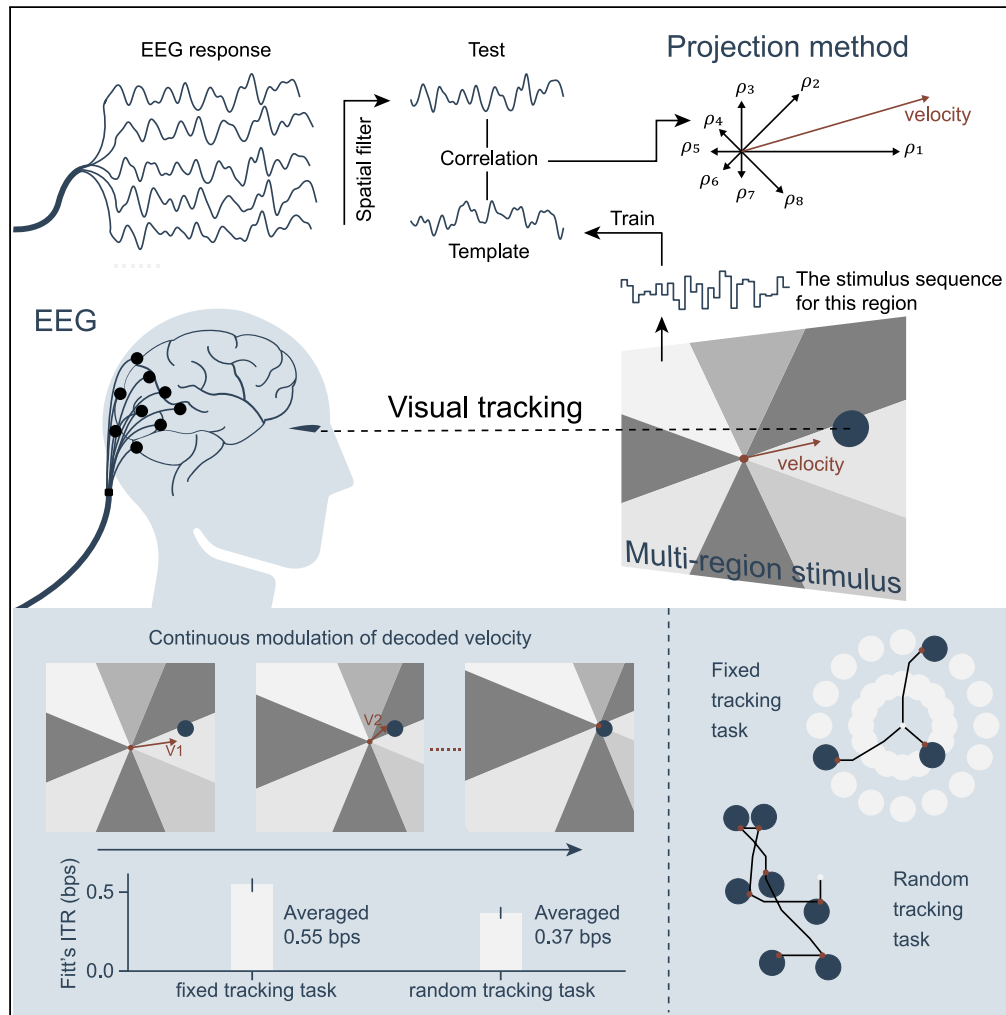## Article
Visual tracking brain-computer interface

Changxing Huang,
Nanlin Shi, Yining
Miao, Xiaogang
Chen, Yijun Wang,
Xiaorong Gao

gxr-dea@mail.tsinghua.edu.cn

Highlights

Advancement in visual BCI
toward natural and efficient
interaction

Novel spatial encoding
stimulus paradigm and
decoding methods for
continuous control

High performance in visual
tracking tasks

Promised real-world
applicability

Article
Visual tracking brain-computer interface

Changxing Huang,[1] Nanlin Shi,[1] Yining Miao,[1] Xiaogang Chen,[2] Yijun Wang,[3] and Xiaorong Gao[1,4,*]

SUMMARY

Brain-computer interfaces (BCIs) offer a way to interact with computers without relying on physical movements. Non-invasive electroencephalography-based visual BCIs, known for efficient speed and calibration ease, face limitations in continuous tasks due to discrete stimulus design and decoding methods. To achieve continuous control, we implemented a novel spatial encoding stimulus paradigm and devised a corresponding projection method to enable continuous modulation of decoded velocity. Subsequently, we conducted experiments involving 17 participants and achieved Fitt's information transfer rate (ITR) of 0.55 bps for the fixed tracking task and 0.37 bps for the random tracking task. The proposed BCI with a high Fitt's ITR was then integrated into two applications, including painting and gaming. In conclusion, this study proposed a visual BCI based-control method to go beyond discrete commands, allowing natural continuous control based on neural activity.

INTRODUCTION

In the age of increasing demand for interaction with electronic devices, brain-computer interfaces (BCIs) offer a new approach that enables computer interaction without relying on muscular activity.[1] Numerous studies have demonstrated the efficacy of BCIs in invasive experimental settings.[2–8] However, non-invasive BCIs based on electroencephalography (EEG) hold the potential for widespread use in everyday life. EEG-based BCIs typically generate commands by decoding motor intention activity or visual evoked responses.[9–12] Among these, visual BCIs, which decode visual responses, have gained popularity due to their efficient communication speed and minimal calibration requirements. They have been successfully applied to various interaction tasks such as typing,[13–16] drawing,[17,18] cursor control,[19–21] and driving robotic devices.[22,23]

Visual BCIs, despite their application in multiple scenarios, are often limited to classification tasks with discrete commands. The classical speller task has achieved an impressive information transfer rate (ITR) of 325 bits per minute (bpm), allowing for the output of one of forty discrete commands in just 0.3 s.[24] However, these discrete commands are not suitable for continuous tasks such as drawing, cursor control, and robotic device control. The main challenge arises from the separation between the visual stimulus and the objects to be controlled. Consequently, users must constantly shift their attention between the visual stimulation and objects like cursor or robotic arms, creating a significant conflict in this setting. For example, in 2018, Chen et al. employed steady state visual evoked potential (SSVEP) based BCI to achieve robotic arm control, relying on step-by-step target selections.[25] However, this approach resulted in a significant time cost during the gaze shift process. In 2010, Muenssinger et al. developed a P300-Brain Painting application, which involved choosing elements and parameters for painting.[17] Similarly, in 2015, Spüler designed a computer interaction system, where mouse movement was achieved through gaze shifts to select directional commands.[21] Both encountered the aforementioned challenge due to the necessity of multiple selections and attention shifts, resulting in inefficient and unnatural human-computer interactions. These studies indicate the inadequacy of discrete decoding methods for achieving continuous control and emphasize the need for innovative solutions within the traditional field of visual BCIs.

To enable natural and efficient interaction, motor imagery BCIs have successfully transformed discrete commands into continuous two-dimensional (2D) cursor control.[26,27] In 2004, Wolpaw et al. demonstrated the control of 2D movement using noninvasive motor imagery BCIs,[28] thereby highlighting that multidimensional control is not exclusive to invasive methods. In 2019, Edelman et al. introduced and validated a noninvasive framework using EEG to achieve the neural control of a robotic device for continuous random target tracking.[27] This work further underscores the capability of non-invasive BCIs to attain continuous control. Compared to the motor imagery paradigm, visual BCIs that emphasize discrete command mode excel in terms of high performance, low calibration effort, and generalizability. Therefore, realizing continuous control for visual BCI is likely to inherit these advantages and ultimately contribute to natural and efficient interaction.

In this study, we introduced a novel visual tracking BCI to achieve continuous cursor control. The pivotal step toward achieving multidimensional continuous control lies in shifting the approach from classifying different classes of patterns to predicting continuous variables from neural activities. By employing the multi-region stimulus paradigm and corresponding projection method, we decoded velocity for cursor movement, achieving a high Fitt's ITR in both fixed and random tracking tasks. The efficiency of the velocity weight used for correcting decoded velocity was demonstrated. We further showcased the versatility of this system through various applications, including painting and
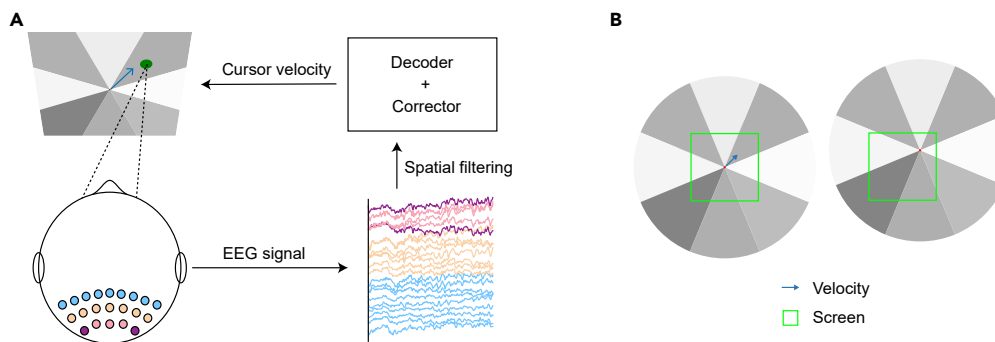
Figure 1. Visual tracking BCI system
(A) Schematic of the visual tracking BCI framework. EEG signals recorded from multiple electrodes are passed to the decoder after spatial filtering. The velocity predicted by the decoder is corrected by the corrector to guide cursor movement.
(B) Stimulus implementation. Cursor movement is achieved by updating and shifting the background image.

gaming. To account for variations among participants, we proposed a pre-test indicator to roughly predict task performance. Finally, we provided a brief analysis of the system's stability and potential avenues for future improvements. Overall, our study represents a significant advancement beyond the limitations of traditional discrete classification modes in visual BCIs, enabling natural and efficient continuous control based on neural activity.

RESULTS

Visual tracking BCI focuses on continuous movement tasks such as cursor control. During the tasks, participants gaze at the intended position under the multi-region stimulus paradigm (see STAR Methods for Stimulus paradigm) while the cursor tracks the position using EEG predictions (see Figure 1A). To enable continuous decoding, we map the correlation coefficients of eight distinct patterns in the paradigm to their corresponding directions (Decoder). Adding velocity weight that was trained using the least square method, a corrected velocity vector can be predicted to control cursor movement (Corrector; see STAR Methods for Decoding algorithm).

Feasibility of projection method

We conducted experiments with 17 participants and observed that the distribution of correlation coefficients (ρ s) reflects the location information of users' visual fixation point (Figure 2C). For the 32 target positions in training Stage II (Figure 2A; see STAR Methods for Experiments materials and design), when the target is located on the center line of the region (referred to as "center situation"), the ρ_i for the corresponding region is significantly larger than the remaining regions (p < 0.001, unpaired t test). Additionally, when comparing the effects of the inner and outer circles, we found that the ρ_i of the outer circle is significantly larger than that of the inner circle (p = 2.85 × 10^{-9}, paired t test), whereas other adjacent ρ s exhibit opposite trends (Figure 2D), indicating ρ s' distribution can effectively align with the target direction and correspond to the target distance. In the case of the target located at the boundary between two adjacent regions (cross situation), both ρ_j and ρ_{j+1} (counter-clockwise ordering) corresponding to the two adjacent regions are relatively large, effectively reflecting the orientation of the target. The differences in ρ s between the inner circle and outer circle exhibit similarities to the center situation (p = 6.77 × 10^{-3} for ρ_j, p = 0.0273 for ρ_{j+1}, paired t test). In summary, the ρ s' distribution can reflect the location information for any target positions. Based on this, a velocity that is able to automatically adjust both magnitude and direction can be derived.

The mechanism through which this method functions is that human vision has receptive field, and the neural activity is jointly encoded by stimulation from all regions in the field.[29] As shown in Figure S3, when fixating on nearby targets (in relation to the cursor), the higher proportion of surrounding stimulus within the field of view exerts a more pronounced influence, causing a dispersed distribution of ρ s toward the neighboring regions (in relation to the target). Conversely, when fixating on distant targets, the reduced proportion of surrounding stimulus within the field of view results in a more centralized distribution of ρ s toward the central regions (refer to Figures 2C and 2D). Here, we have generated the distribution of ρ s which can represent the users' fixation location by employing the multi-region stimulus paradigm. By subsequently incorporating the corresponding decoding algorithm, we are able to generate the desired continuous variable, which is the velocity.

Effect of velocity prediction and velocity weight

To assess the performance of the velocity prediction and the effect of the velocity weight, we employed angular error and vector error as evaluation metrics (see STAR Methods for BCI performance evaluation). We found that the decoding algorithm can effectively estimate the intended velocity, and the utilization of corrected velocity weight significantly minimizes both angular error and vector error. Figure 3A illustrates the decoded velocity of all participants in three typical target positions. All velocities are able to guide the cursor to approach the target, and in the case of certain subjects, the velocity accurately points to the target with the appropriate size and direction. In order to observe the characteristics of velocity weight, we presented the average corrected velocity weight for all participants corresponding to
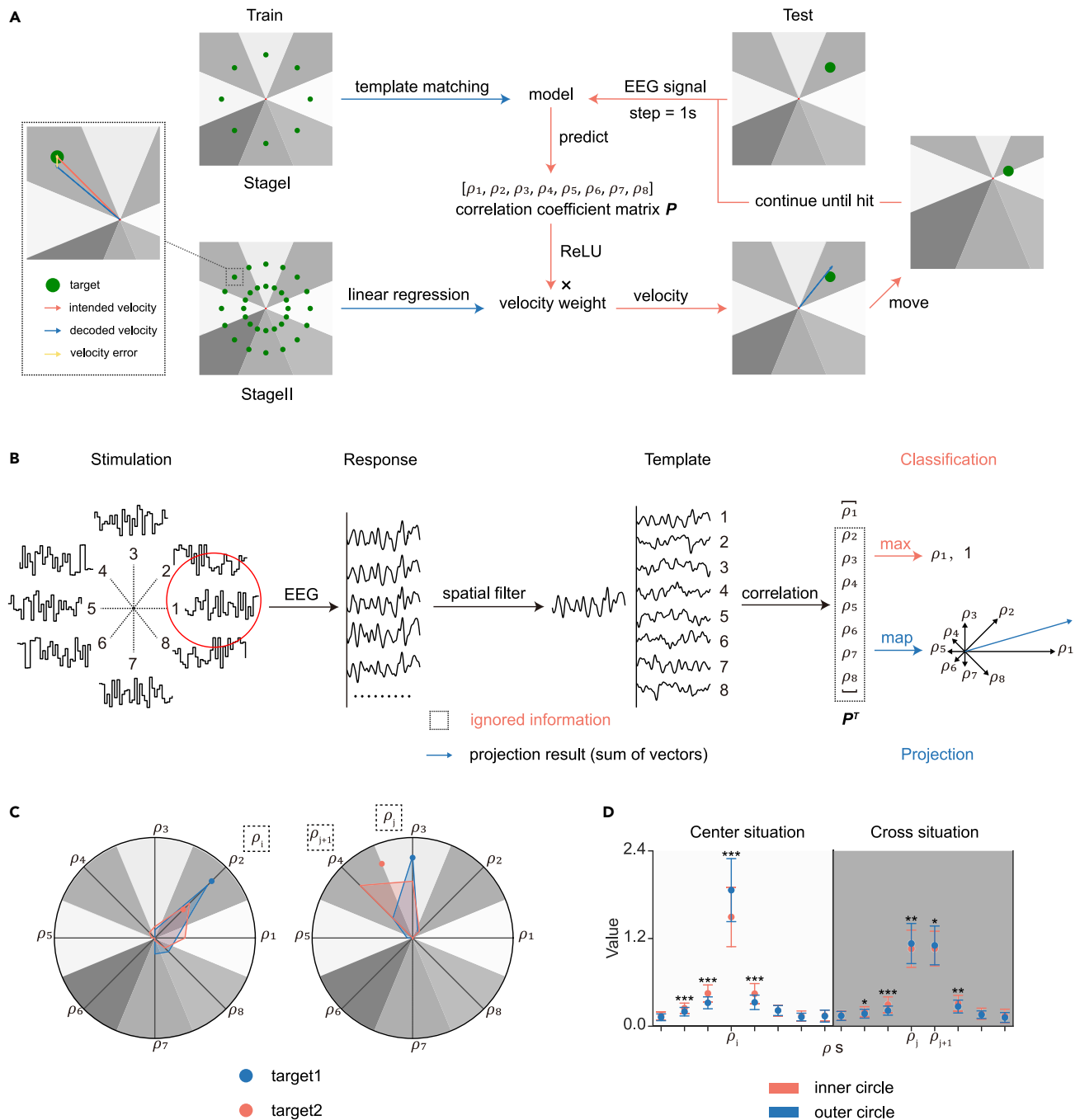
Figure 2. Decoder design and feasibility

(A) The main experiment consists of training and testing parts. The training part is divided into stage I (for training template matching model) and stage II (for training velocity weight by linear regression). In the testing part, cursor moves at the decoded velocity until it hits the target.

(B) Diagram of the decoding method. Projection method projects correlation coefficients onto the corresponding directions, utilizing information that is ignored by the classification method.

(C) Radar plot of the correlation coefficient distribution. Left: the examples of distribution for inner circle and outer circle positions; right: the examples of distribution for center and cross situations.

(D) Comparison of ρ s' value in different conditions. Bars indicate mean $\pm$ SD. Paired t test ($N = 17$): *p < 0.05, **p < 0.01, ***p < 0.001.
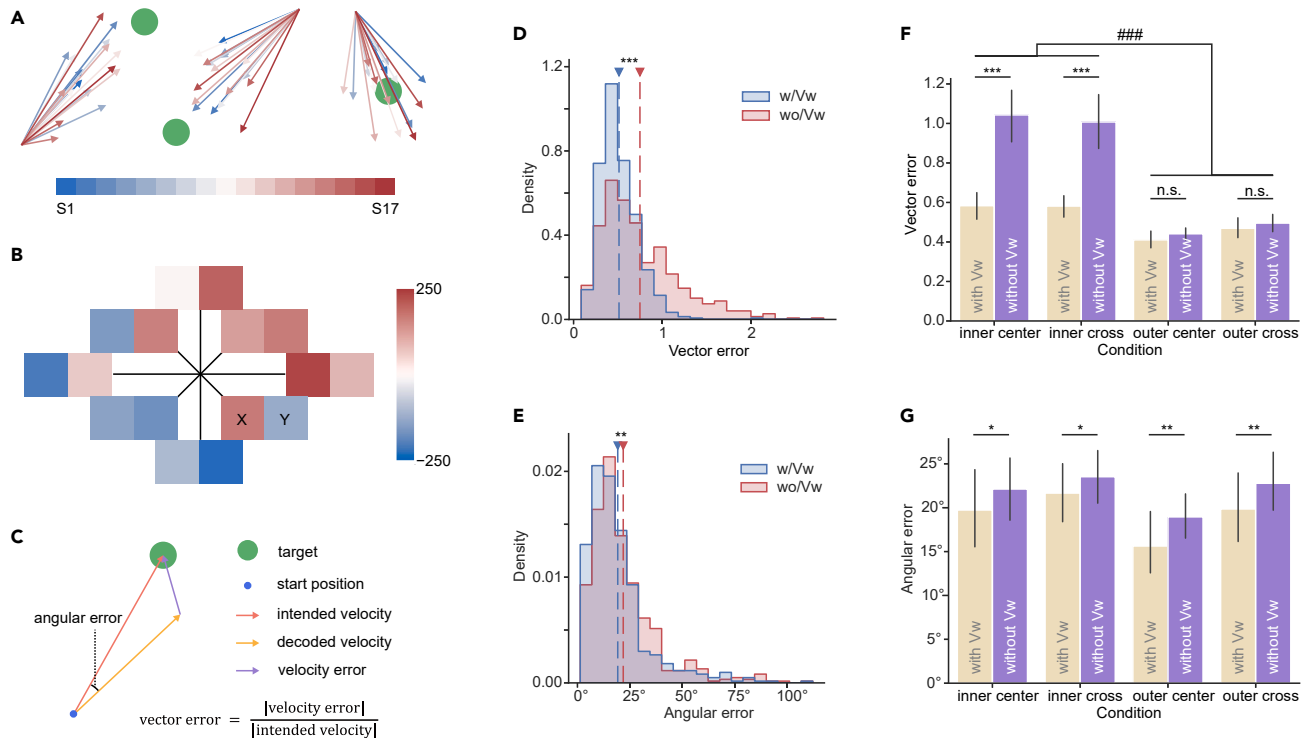
Figure 3. Velocity prediction and velocity weight

(A) Decoded velocity of all participants in three typical target positions.

(B) Velocity weight averaged for all participants ($N = 17$), conforming to the form of the Cartesian coordinate system.

(C) Schematic diagram of vector error and angular error.

(D) Distribution of vector error for all participants using corrected velocity weight (w/Vw) and using only the initial velocity weight (wo/Vw) (paired t test: p = 4.396×10^{-5}, $N = 17$; w/Vw: mean = 0.51, SD = 0.08; wo/Vw: mean = 0.75, SD = 0.14).

(E) Distribution of angular error for all participants (paired t test: p = 0.0016, $N = 17$; w/Vw: mean = 19.1°, SD = 7.2°; wo/Vw: mean = 21.7°, SD = 6.0°).

(F) Vector error of different conditions ("inner center": at the inner circle and the center of the region, "inner cross": at the inner circle and the boundary between the regions, so as for "outer center" and "outer cross").

(G) Angular error of different conditions. Bars indicate mean ± SD. Statistical analysis using two-way ANOVA ((F), $N = 34$ per group, ###p < 0.001) and paired t test ((D and F), $N = 17$, *p < 0.5, **p < 0.01, ***p < 0.001).

the projection direction in Figure 3B. As anticipated, the positive and negative relationship of velocity weight conforms to the Cartesian co-ordinate system, suggesting that the decoding algorithm leverages spatial encoding information from the stimulus paradigm. Furthermore, the selected stimulus sequences exhibit consistency within the population with no significant directional deviation (aligning with the charac-teristics of the white noise [WN] sequence). The velocity weight for each participant was displayed in Figure S4.

We then compared the vector error and angular error of the decoded velocity between using the corrected velocity weight (w/Vw) and using only the initial velocity weight (wo/Vw), and found that the use of corrected velocity weight significantly reduced the vector error (Fig-ure 3D; paired t test: p = 4.396×10^{-5}, $N = 17$; w/Vw: mean = 0.51, SD = 0.08; wo/Vw: mean = 0.75, SD = 0.14) and angular error (Figure 3E; paired t test: p = 0.0016, $N = 17$; w/Vw: mean = 19.1°, SD = 7.2°; wo/Vw: mean = 21.7°, SD = 6.0°) of the decoded velocity. Moreover, we examined the impact of target location (inner/outer, cross/center) on vector error and angular error. A significant improvement of vector error was noted in "inner center" and "inner cross" conditions (Figure 3F; p < 0.001, paired t test), while no significant improvement was noted for "outer center" and "outer cross" conditions. This demonstrated that the corrected velocity weight mainly rectified the vector error of the inner circle position, which aligns with the finding that the vector error of the inner circle position was significantly larger than that of the outer circle position ($F = 102.2$, p = 3.53×10^{-18}, two-way ANOVA). In terms of angular error, a significant improvement was observed in all conditions (Figure 3G; "inner center": p < 0.05, "inner cross": p < 0.05, "outer center": p < 0.01, "outer cross": p < 0.01). In summary, we discovered that the corrected velocity weight can significantly reduce both angular error and vector error. Particularly, the impact on vector error is most pro-nounced in the inner circle situation.

Tracking task performance

Based on the training session, we successfully achieved visual tracking in both fixed and random tracking tasks (see STAR Methods for Ex-periments materials and design). Figure 4A illustrates the tracking trajectories of the fixed tracking task and random tracking task (example
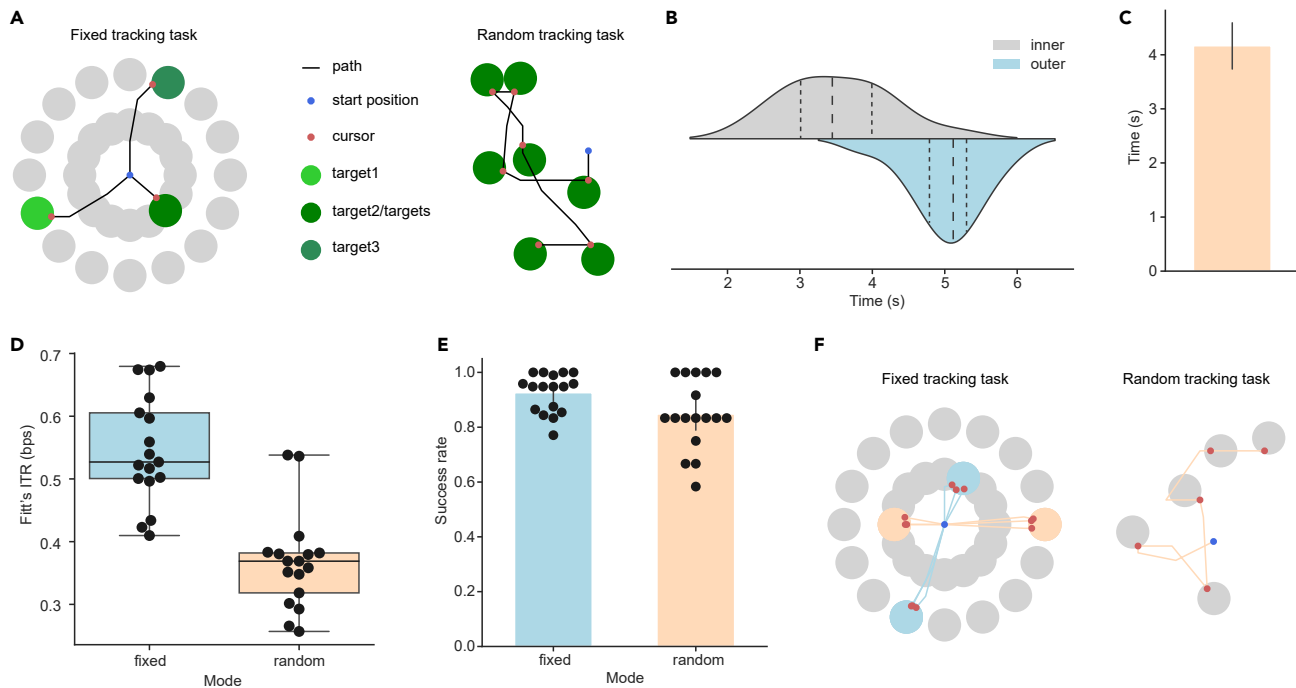
Figure 4. Visual tracking task performance

(A) Schematic of fixed and random tracking tasks.

(B) Distribution of time to target for fixed tracking task in inner circle and outer circle conditions. Dashed lines are quarters.

(C) Time to target for the random tracking task. Bars indicate mean ± 95% CI.

(D) Fitt's ITR of fixed and random tracking tasks (N = 17, medians, interquartile ranges, maxima, and minima).

(E) Success rate of fixed and random tracking tasks (N = 17, bars indicate mean ± 95% CI).

(F) Tracking trajectories of fixed and random tracking tasks (four typical positions for the fixed tracking task, and the first 5 targets for the random tracking task).

results come from Subject 13 and Subject 1, see Table S1). To evaluate the performance of tracking tasks, we calculated metrics such as time to target, success rate, and Fitt's ITR (see STAR Methods for BCI performance evaluation). Figure 4B shows the distribution of time to target of the fixed tracking task for all participants. We separated the conditions into "inner" (inner circle) and "outer" (outer circle), and as expected, time to target of the inner circle was smaller than that of the outer circle (inner: mean = 3.53, SD = 0.74; outer: mean = 5.03, SD = 0.51). The average time to target of the random tracking task was 4.17 s, with an SD of 0.93 (as in Figure 4C). We then obtained Fitt's ITR of the fixed tracking task and random tracking task for each participant (Figure 4D). The average of Fitt's ITR in the fixed tracking task was 0.55 bps with a maximum of 0.68 bps (SD = 0.09), and the average of Fitt's ITR in the random tracking task was 0.37 bps with a maximum of 0.54 bps (SD = 0.08). The results are comparable to related invasive studies, which typically yield a Fitt's ITR of 0.4 bps.[26] We attributed the lower Fitt's ITR in the random tracking task to the fact that the distance to the random target for successful trials was close to the start position, while for farther targets, trials were more prone to failure. Figure 4E illustrates the success rate for the two tasks (fixed: mean = 92.6%, SD = 7.1%; random: mean = 84.8%, SD = 12.9%). Tracking trajectories for all participants in both tasks are depicted in Figures S5 and S6. And here, we showed the tracking trajectories of the representative subject for four specific target positions of fixed tracking task, including three tracking paths for each target position (Subject 3; Figure 4F, left). Additionally, we displayed trajectories for the first 5 target positions of the random tracking task (Figure 4F, right). For both the fixed and random tracking tasks, we recorded the videos to show the process (see Videos S1 and S2). And Video S5 shows the experimental environment, where participant was conducting random tracking tasks. The aforementioned results demonstrate our successful achievement of visual tracking and the attainment of a level of cursor movement control that matches invasive methods.

Applications

Painting

Expanding on the random tracking task, we designed a painting application for all participants to experience after the testing part. We positioned two windows on the screen: one for stimulation and the other for painting display (see Figure 5A). During the painting experience, we allowed participants to freely move their gaze to experience controlling the brush. The results were collected as a questionnaire asking participants, "Do you think the brush draws according to your will?" (Almost positive), and for participants who experienced BCI speller, we also asked, "Is the visual load of this application larger than BCI speller?" (All are negative; see Table S1). Figure 5B shows the painting of one participant (Subject 3), depicting the image of a dog. The painting process for writing number 2 can be seen in Video S3.
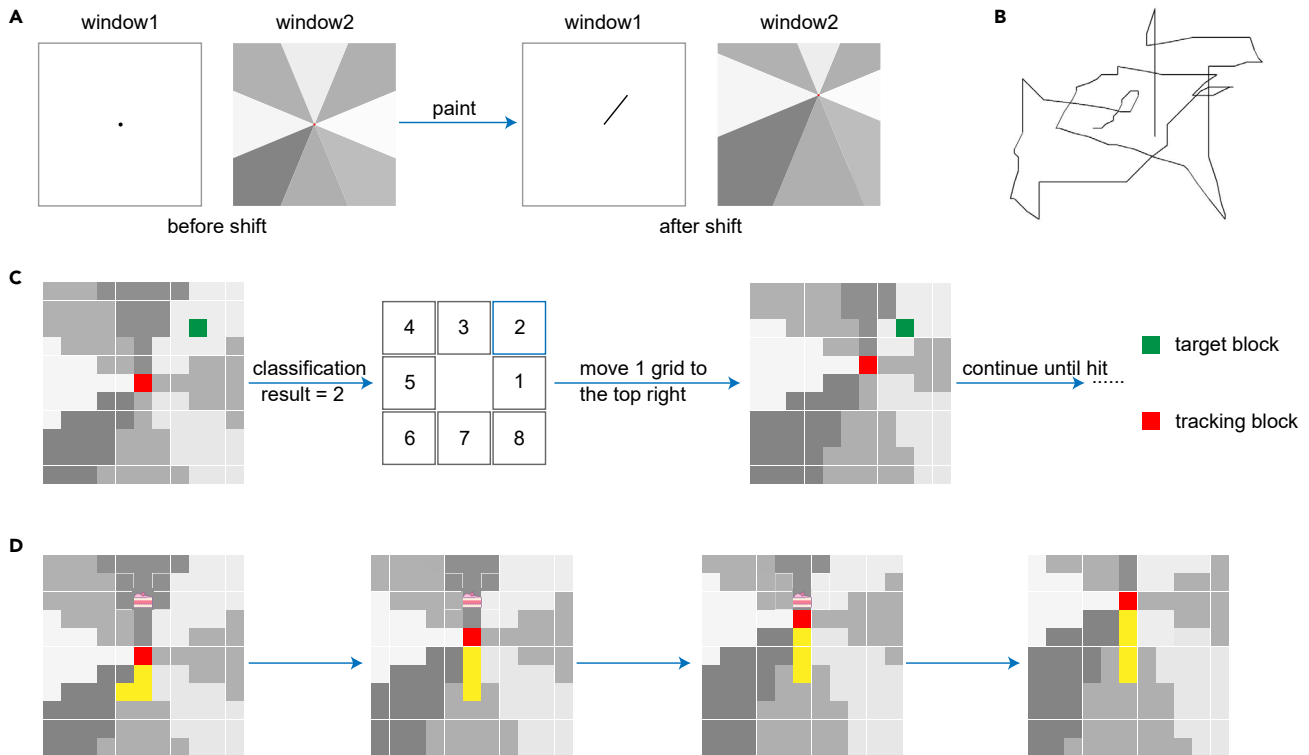
Figure 5. Applications of visual tracking BCI

(A) Design of painting application. Window 1 for painting display, and window 2 for stimulation.

(B) The painting result of one participant (Subject 3) depicts a dog.

(C) Schematic diagram of the visual tracking brain-computer interface in discrete cases.

(D) Design and process of the snake game.

Snake game

To facilitate control for the snake game, we simplified the visual tracking BCI to a discrete case. Firstly, the stimulus region consists of stimulus blocks arranged radially around the central red tracking block. Secondly, the decoding algorithm only uses traditional classification methods. When the user fixates on the green target block (see Figure 5C), the system computes the classification result corresponding to the region where the target is being fixated. Subsequently, the red tracking block moves one grid space toward that region. This process continues until the red tracking block reaches and "hits" the green target block. We used the discrete visual tracking BCI technology to design the snake game (as shown in Figure 5D; red block: head; yellow block: body). In the game, users control the snake's movement up, down, left, and right by fixating on the corresponding regions. The regions in other directions do not possess movement control functions, allowing users to shift their gaze to observe the food's position without triggering unnecessary movement of the tracking block. The users need to control the snake to obtain food and avoid hitting the snake's body or the edge of the screen. At the end, a score that stands for the length of the snake will be given. 17 participants played the snake game and got their scores (see Figure S7; maxima = 14). The gaming process was recorded in Video S4.

DISCUSSION

In this study, we proposed a novel spatial encoding stimulus paradigm and the corresponding velocity projection method to achieve continuous visual tracking based on noninvasive EEG. We achieved ideal results in velocity prediction and cursor control tasks across 17 subjects and successfully implemented various applications, showcasing the efficiency and practical utility of the visual tracking BCI. Furthermore, we have elucidated the mechanism of the stimulus paradigm through the lens of the receptive field, which signifies a novel approach to continuous decoding. In summary, this study transcends traditional discrete visual BCI and achieves continuous control based on a noninvasive method.

Interpretability of BCI performance

As we can notice, the continuous control performance experiences inter-subject variability. We have identified a metric for pre-estimating the performance of the participants in tracking tasks. The metric was defined as the length of the projection result from the data of training Stage
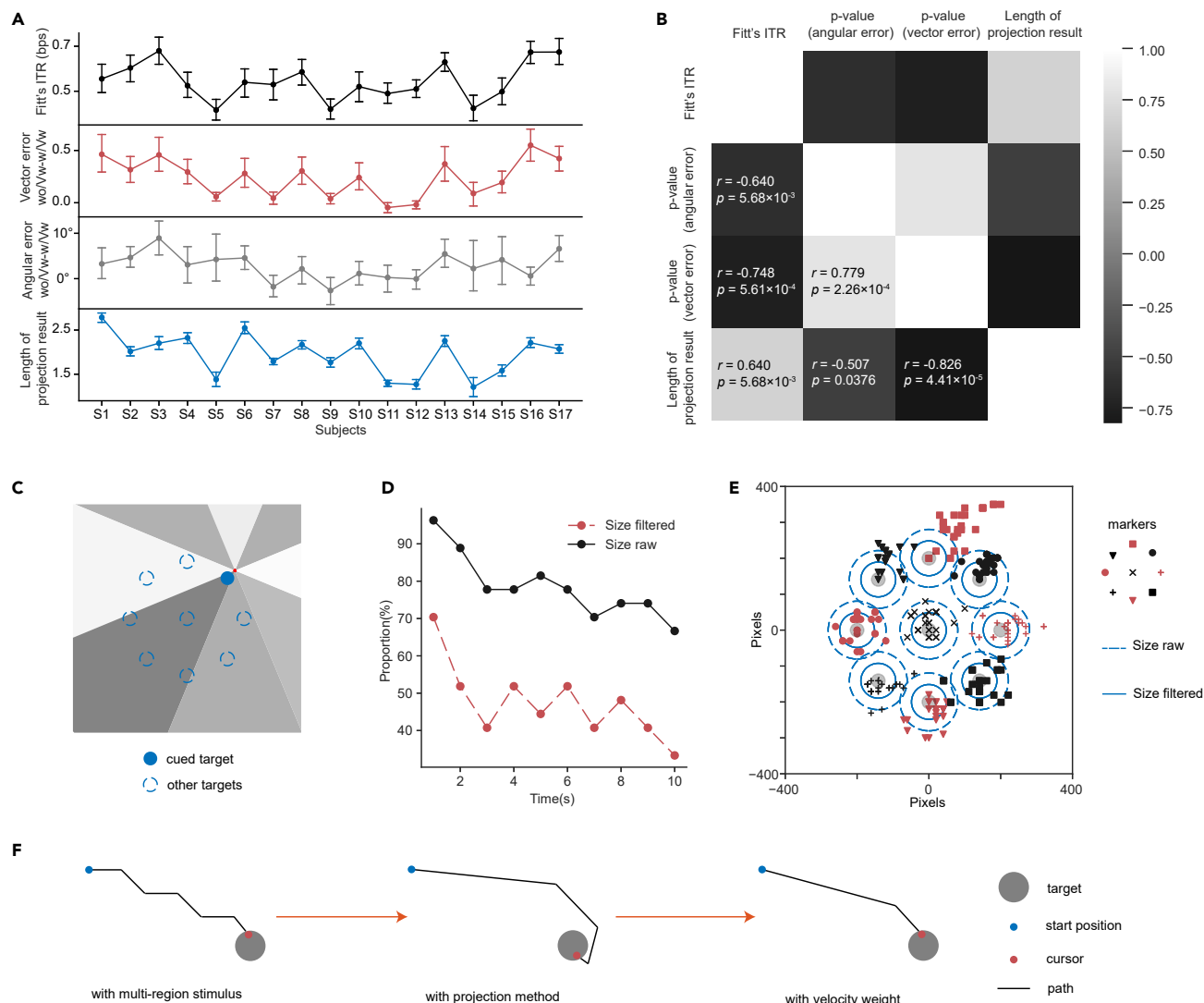
Figure 6. Further evaluation and examination of the visual tracking BCI

(A) Fitt's ITR, vector error difference (wo/Vw-w/Vw), angular error difference (wo/Vw-w/Vw), and length of projection result across all participants. Bars indicate mean ± 95% CI.

(B) Spearman correlation coefficient matrix between Fitt's ITR, p value of vector error difference, p value of angular error difference, and length of projection result (N = 17).

(C) Schematic of the jitter inspection experiment.

(D) Proportion of jitter positions within "size filtered" and within "size raw" over time.

(E) Map of jitter positions for each target. Different markers correspond to different targets.

(F) Schematic of path optimization.

II, which exhibited a robust correlation with other metrics of BCI performance. Figure 6A illustrates Fitt's ITR, vector error difference, angular error difference, and the length of the projection result across participants, revealing a distinct correlation among them. Additionally, we assessed the significance levels (p values) of vector error difference and angular error difference to replace the two differences and computed the Spearman correlation coefficients among the four metrics (see Figure 6B). The result shows a positive correlation between the length of the projection result and Fitt's ITR ($r = 0.640$, p = 5.68×10^{-3}), suggesting that it can serve as an estimator for the task performance level of the participants. It is worth noting that the length of the projection result exhibits a negative correlation with both the p value of vector error ($r = -0.826$, p = 4.41×10^{-5}) and angular error ($r = -0.507$, p = 0.0376), which illustrates that the length of projection result can predict the effectiveness of the corrected velocity weight. Based on our current observations, two situations lead to a longer projection result, both rooted in the subject's heightened responsiveness to stimulus encoding. Firstly, when the subject's attention field of vision is more focused, it is less influenced by the surrounding stimulus regions. This results in the subject primarily receiving stimulation from the region

within the visual field, making the projection vector corresponding to that region dominant. Secondly, subjects with a broader field of vision are stimulated by more regions. However, owing to the spatial encoding characteristics, the sum of vectors (projection result) accumulates in the target direction and nullifies in other directions. This cumulative effect ultimately yields a longer projection result.

Gaze fixation and path optimization

For human-computer interaction similar to cursor control, in addition to the requirement of reaching the target, the accuracy of the movement path largely determines the efficiency and functionality of the system. As shown in Figure 6F, path optimization is essential for the performance of visual tracking BCI. Utilizing solely the multi-region stimulation paradigm can only facilitate discrete visual tracking. However, by incorporating the projection method, continuous visual tracking can be achieved. Furthermore, in combination with velocity weight, velocity can be refined for enhanced accuracy and reduced tracking time. However, besides achieving continuous smooth tracking, knowing when to stop is also crucial to enabling the functionality of selection. Therefore, we investigated the issue of cursor jitter to determine the range and time of the users' gaze fixation. To detect the jitter of the tracking point (cursor) in the visual tracking BCI, we conducted an experiment in which users were instructed to fixate on the cued target (including 9 target positions; Figure 6C) while recording the jitter positions within the subsequent 10 s. Subject 6 participated in the experiment, which was repeated 3 times for each target position. For the evaluation of jitter intensity, we employed the target recognition range from eye-tracking technology as a reference.[30] This range measured 1.9 × 2.35 cm (size filtered; square diagonal length as circular diameter) when utilizing a filter and 3.28 × 3.78 cm (size raw) when not. Then, we found that this technology can achieve the maintenance of gaze fixation within a certain time and range. Figure 6D shows the variation of the proportion of jitter positions within "size filtered" and within "size raw" over time (size raw: 1s, 96.3%; size filtered: 1s, 70.4%). And we presented all jitter positions for each target in Figure 6E. In the context of the tracking task, we define "hit" as the moment that the cursor makes contact with the target. However, during that instant, the cursor's velocity is not zero due to the 1-s time step employed. As a result, we ascertained whether the cursor continues to maintain "hit" the target after a certain duration (see Figure S8; 0.2 s, 83.5%).

Comparison with other BCIs

There are currently two established BCI methods for cursor movement tasks. One involves an invasive spike-based BCI that utilizes Kalman filter-based algorithms.[26,31,32] The other is a non-invasive motor imagery BCI, primarily employing EEG signals and the velocity algorithm proposed by Wolaw et al.[28] In our study, we achieved an average Fitt's ITR of 0.55 bps in the fixed visual tracking task, often referred to as a center-out task in other research. These results compare favorably to some spike-based BCIs, which have reported Fitt's ITR of 0.4 bps,[26] although they fall below the highest recorded rates for individuals ranging from 1.4 to 4.2 bps.[2] In comparison to non-invasive BCIs, our results surpass the 0.29 bps calculated from the presented data of a 2D BCI task using Wolpaw's method.[33]

BCI spellers typically require larger target sizes to achieve optimal performance,[34] constrained by the discrete decoding method. As the number of targets increases, the screen size also expands.[16,35] This constraint hinders the applicability of BCI spellers in complex scenarios and lacks suitability for smaller devices like smartphones. We posit that visual tracking BCI can achieve high performance even with a smaller screen size and reduced stimulus contrast, making it well-suited to address these challenges in alternative devices and complex scenes. Visual tracking BCI also holds the potential for further reduction in visual fatigue. Illustrated in Figure S9, we propose an alternative stimulus paradigm that incorporates peripheral visual field stimulation. In this paradigm, velocity decoding is executed, with no stimulation occurring in the central area, offering a novel and refreshing visual experience. This approach suggests that exploring similar paradigms could provide additional avenues for addressing the visual load challenges encountered by visual BCIs. In the Applications section, we outlined the potential scenarios for applying visual tracking BCI. It is evident that this technology can readily adapt to diverse tasks for developing continuous movement control, and there is the prospect that visual writing through BCIs may become feasible in the future. As illustrated in Figure S9, we also anticipate the utilization of visual tracking BCI in virtual reality scenarios.

Limitations of the study

This research has some limitations and we believe that future improvements should focus on stimulus design and algorithm optimization, real-time asynchronous system construction, and click function implementation. The dynamic sequences we employed here are the WN sequences; however, it's worth noting that other sequences, such as sinusoidal sequences or m-sequences,[36] can also be tested. Importantly, stimulus sequences from different regions may not necessarily have an equal impact on participants.[37] This can pose challenges when tracking targets within specific directional ranges. In this study, we did not extensively investigate the varying effects of different stimulus sequences on participants. However, exploring these differences could potentially provide a path to enhance BCI performance.[38] From a different perspective, non-equivalent stimulus sequences, such as sinusoidal sequences with different frequencies, may hold greater potential in certain scenarios with preferences for specific directions. In this study, we employed task-related component analysis (TRCA) as the template matching method. However, other classification-based algorithms such as task-discriminant component analysis (TDCA),[39] or filter bank canonical correlation analysis (FBCCA)[13] for SSVEP are also possible. Furthermore, it's important to note that we utilized only one type of target size, and the training of velocity weight relied solely on a linear method. Consequently, we believe that there is substantial potential for enhancing the task performance of visual tracking. We aspire for future work to employ shorter time steps or build asynchronous systems[40] to achieve real-time visual tracking. Alternatively, a sliding window method can be used to collect 1 s of data every 0.1 s for analysis to improve the temporal resolution of decoding. For the click function, we aim to realize it through

dwell time[41] or in combination with the BCI of motor imagery.[42] Combining these improvements, this technology will ultimately achieve a more efficient, comfortable, and natural human-computer interaction effect.

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
  - Stimulus paradigm
  - Experiments materials and design
  - Decoding algorithms
  - BCI performance evaluation
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.109376.

AUTHOR CONTRIBUTIONS

Conceptualization: C.H., N.S., and X.G.; methodology: C.H.; investigation: C.H. and Y.M.; visualization: C.H. and N.S.; supervision: X.C., Y.W., and X.G.; data collection and curation: C.H. and Y.M.; data processing: C.H.; Writing—original draft: C.H., N.S., X.C., and Y.W.; writing—review and editing: all authors.

DECLARATION OF INTERESTS

The authors declare no competing interest.

REFERENCES

1. Gao, X., Wang, Y., Chen, X., and Gao, S. (2021). Interface, interaction, and intelligence in generalized brain–computer interfaces. Trends Cogn. Sci. 25, 671–684. https://doi.org/10.1016/j.tics.2021.04.003.

2. Pandarinath, C., Nuyujukian, P., Blabe, C.H., Sorice, B.L., Saab, J., Willett, F.R., Hochberg, L.R., Shenoy, K.V., and Henderson, J.M. (2017). High performance communication by people with paralysis using an intracortical brain-computer interface. Elife 6, e18554. https://doi.org/10.7554/eLife.18554.

3. Degenhart, A.D., Bishop, W.E., Oby, E.R., Tyler-Kabara, E.C., Chase, S.M., Batista, A.P., and Yu, B.M. (2020). Stabilization of a brain–computer interface via the alignment of low-dimensional spaces of neural activity. Nat. Biomed. Eng. 4, 672–685. https://doi.org/10.1038/s41551-020-0542-9.

4. Willett, F.R., Avansino, D.T., Hochberg, L.R., Henderson, J.M., and Shenoy, K.V. (2021). High-performance brain-to-text communication via handwriting. Nature 593, 249–254. https://doi.org/10.1038/s41586-021-03506-2.

5. Anumanchipalli, G.K., Chartier, J., and Chang, E.F. (2019). Speech synthesis from neural decoding of spoken sentences. Nature 568, 493–498. https://doi.org/10.1038/s41586-019-1119-1.

6. Metzger, S.L., Littlejohn, K.T., Silva, A.B., Moses, D.A., Seaton, M.P., Wang, R., Dougherty, M.E., Liu, J.R., Wu, P., Berger, M.A., et al. (2023). A high-performance neuroprosthesis for speech decoding and avatar control. Nature 620, 1037–1046. https://doi.org/10.1038/s41586-023-06443-4.

7. Willett, F.R., Kunz, E.M., Fan, C., Avansino, D.T., Wilson, G.H., Choi, E.Y., Kamdar, F., Glasser, M.F., Hochberg, L.R., Druckmann, S., et al. (2023). A high-performance speech neuroprosthesis. Nature 620, 1031–1036. https://doi.org/10.1038/s41586-023-06377-x.

8. Hochberg, L.R., Serruya, M.D., Friehs, G.M., Mukand, J.A., Saleh, M., Caplan, A.H., Branner, A., Chen, D., Penn, R.D., and Donoghue, J.P. (2006). Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature 442, 164–171. https://doi.org/10.1038/nature04970.

9. Neuper, C., Scherer, R., Reiner, M., and Pfurtscheller, G. (2005). Imagery of motor actions: Differential effects of kinesthetic and visual–motor mode of imagery in single-trial EEG. Cognit. Brain Res. 25, 668–677. https://doi.org/10.1016/j.cogbrainres.2005.08.014.

10. Pfurtscheller, G., and Neuper, C. (2001). Motor imagery and direct brain-computer communication. Proc. IEEE 89, 1123–1134. https://doi.org/10.1109/5.939829.

11. Sutter, E.E. (1992). The brain response interface: communication through visually-induced electrical brain responses. J. Microcomput. Appl. 15, 31–45. https://doi.org/10.1016/0745-7138(92)90045-7.

12. Bin, G., Gao, X., Wang, Y., Hong, B., and Gao, S. (2009). VEP-based brain-computer interfaces: time, frequency, and code modulations [Research Frontier]. IEEE Comput. Intell. Mag. *4*, 22–26. https://doi.org/10.1109/MCI.2009.934562.

13. Chen, X., Wang, Y., Nakanishi, M., Gao, X., Jung, T.-P., and Gao, S. (2015). High-speed spelling with a noninvasive brain–computer interface. Proc. Natl. Acad. Sci. USA *112*, E6058–E6067. https://doi.org/10.1073/pnas.1508080112.

14. Lin, Z., Zhang, C., Wu, W., and Gao, X. (2006). Frequency Recognition Based on Canonical Correlation Analysis for SSVEP-Based BCIs. IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. *53*, 2610–2614. https://doi.org/10.1109/TBME.2006.886577.

15. Bin, G., Gao, X., Yan, Z., Hong, B., and Gao, S. (2009). An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method. J. Neural Eng. *6*, 046002. https://doi.org/10.1088/1741-2560/6/4/046002.

16. Chen, Y., Yang, C., Ye, X., Chen, X., Wang, Y., and Gao, X. (2021). Implementing a calibration-free SSVEP-based BCI system with 160 targets. J. Neural Eng. *18*, 046094. https://doi.org/10.1088/1741-2552/ac0bfa.

17. Muenssinger, J.I., Halder, S., Kleih, S., Furdea, A., Raco, V., Hoesle, A., and Kubler, A. (2010). Brain Painting: First Evaluation of a New Brain–Computer Interface Application with ALS-Patients and Healthy Volunteers. Front. Neurosci. *4*, 182.

18. Zickler, C., Halder, S., Kleih, S.C., Herbert, C., and Kübler, A. (2013). Brain Painting: Usability testing according to the user-centered design in end users with severe motor paralysis. Artif. Intell. Med. *59*, 99–110. https://doi.org/10.1016/j.artmed.2013.08.003.

19. Bi, L., Lian, J., Jie, K., Lai, R., and Liu, Y. (2014). A speed and direction-based cursor control system with P300 and SSVEP. Biomed. Signal Process. Control *14*, 126–133. https://doi.org/10.1016/j.bspc.2014.07.009.

20. Allison, B.Z., Brunner, C., Altstätter, C., Wagner, I.C., Grissmann, S., and Neuper, C. (2012). A hybrid ERD/SSVEP BCI for continuous simultaneous two dimensional cursor control. J. Neurosci. Methods *209*, 299–307. https://doi.org/10.1016/j.jneumeth.2012.06.022.

21. Spüler, M. (2015). A Brain-Computer Interface (BCI) system to use arbitrary Windows applications by directly controlling mouse and keyboard. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1087–1090. https://doi.org/10.1109/EMBC.2015.7318554.

22. Kishore, S., González-Franco, M., Hintemüller, C., Kapeller, C., Guger, C., Slater, M., and Blom, K.J. (2014). Comparison of SSVEP BCI and Eye Tracking for Controlling a Humanoid Robot in a Social Environment. Presence. (Camb). *23*, 242–252. https://doi.org/10.1162/PRES_a_00192.

23. Müller-Putz, G.R., and Pfurtscheller, G. (2008). Control of an Electrical Prosthesis With an SSVEP-Based BCI. IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. *55*, 361–364. https://doi.org/10.1109/TBME.2007.897815.

24. 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 (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. *65*, 104–112. https://doi.org/10.1109/TBME.2017.2694818.

25. Chen, X., Zhao, B., Wang, Y., Xu, S., and Gao, X. (2018). Control of a 7-DOF Robotic Arm System With an SSVEP-Based BCI. Int. J. Neur. Syst. *28*, 1850018. https://doi.org/10.1142/S0129065718500181.

26. Silversmith, D.B., Abiri, R., Hardy, N.F., Natraj, N., Tu-Chan, A., Chang, E.F., and Ganguly, K. (2021). Plug-and-play control of a brain–computer interface through neural map stabilization. Nat. Biotechnol. *39*, 326–335. https://doi.org/10.1038/s41587-020-0662-5.

27. Edelman, B.J., Meng, J., Suma, D., Zurn, C., Nagarajan, E., Baxter, B.S., Cline, C.C., and He, B. (2019). Noninvasive neuroimaging enhances continuous neural tracking for robotic device control. Sci. Robot. *4*, eaaw6844. https://doi.org/10.1126/scirobotics.aaw6844.

28. Wolpaw, J.R., and McFarland, D.J. (2004). Control of a two-dimensional movement signal by a noninvasive brain-computer interface in humans. Proc. Natl. Acad. Sci. USA *101*, 17849–17854. https://doi.org/10.1073/pnas.0403504101.

29. Wandell, B.A., and Winawer, J. (2015). Computational neuroimaging and population receptive fields. Trends Cogn. Sci. *19*, 349–357. https://doi.org/10.1016/j.tics.2015.03.009.

30. Feit, A., Williams, S., Toledo, A., Paradiso, A., Kulkarni, H., Kane, S., and Morris, M. (2017). Toward Everyday Gaze Input: Accuracy and Precision of Eye Tracking and Implications for Design. In Proceedings of the 2017 Chi conference on human factors in computing systems, pp. 1118–1130. https://doi.org/10.1145/3025453.3025599.

31. Gilja, V., Nuyujukian, P., Chestek, C.A., Cunningham, J.P., Yu, B.M., Fan, J.M., Churchland, M.M., Kaufman, M.T., Kao, J.C., Ryu, S.I., and Shenoy, K.V. (2012). A high-performance neural prosthesis enabled by control algorithm design. Nat. Neurosci. *15*, 1752–1757. https://doi.org/10.1038/nn.3265.

32. Hochberg, L.R., Bacher, D., Jarosiewicz, B., Masse, N.Y., Simeral, J.D., Vogel, J., Haddadin, S., Liu, J., Cash, S.S., van der Smagt, P., and Donoghue, J.P. (2012). Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. Nature *485*, 372–375. https://doi.org/10.1038/nature11076.

33. Felton, E.A., Radwin, R.G., Wilson, J.A., and Williams, J.C. (2009). Evaluation of a modified Fitts law brain–computer interface target acquisition task in able and motor disabled

individuals. J. Neural Eng. *6*, 056002. https://doi.org/10.1088/1741-2560/6/5/056002.

34. Duszyk, A., Bierzyńska, M., Radzikowska, Z., Milanowski, P., Kuś, R., Suffczyński, P., Michalska, M., Łabęcki, M., Zwoliński, P., and Durka, P. (2014). Towards an Optimization of Stimulus Parameters for Brain-Computer Interfaces Based on Steady State Visual Evoked Potentials. PLoS One *9*, e112099. https://doi.org/10.1371/journal.pone.0112099.

35. Chen, X., Liu, B., Wang, Y., and Gao, X. (2022). A Spectrally-Dense Encoding Method for Designing a High-Speed SSVEP-BCI With 120 Stimuli. IEEE Trans. Neural Syst. Rehabil. Eng. *30*, 2764–2772. https://doi.org/10.1109/TNSRE.2022.3208717.

36. Bin, G., Gao, X., Wang, Y., Li, Y., Hong, B., and Gao, S. (2011). A high-speed BCI based on code modulation VEP. J. Neural Eng. *8*, 025015. https://doi.org/10.1088/1741-2560/8/2/025015.

37. Gulbinaite, R., Roozendaal, D.H.M., and VanRullen, R. (2019). Attention differentially modulates the amplitude of resonance frequencies in the visual cortex. Neuroimage *203*, 116146. https://doi.org/10.1016/j.neuroimage.2019.116146.

38. Ye, X., Yang, C., Chen, Y., Wang, Y., Gao, X., and Zhang, H. (2022). Multisymbol Time Division Coding for High-Frequency Steady-State Visual Evoked Potential-Based Brain-Computer Interface. IEEE Trans. Neural Syst. Rehabil. Eng. *30*, 1693–1704. https://doi.org/10.1109/TNSRE.2022.3183087.

39. Liu, B., Chen, X., Shi, N., Wang, Y., Gao, S., and Gao, X. (2021). Improving the Performance of Individually Calibrated SSVEP-BCI by Task-Discriminant Component Analysis. IEEE Trans. Neural Syst. Rehabil. Eng. *29*, 1998–2007. https://doi.org/10.1109/TNSRE.2021.3114340.

40. Chen, Y., Sun, X., and Jin, Y. (2020). Communication-Efficient Federated Deep Learning With Layerwise Asynchronous Model Update and Temporally Weighted Aggregation. IEEE Trans. Neural Netw. Learn. Syst. *31*, 4229–4238. https://doi.org/10.1109/TNNLS.2019.2953131.

41. Carter, B.T., and Luke, S.G. (2020). Best practices in eye tracking research. Int. J. Psychophysiol. *155*, 49–62. https://doi.org/10.1016/j.ijpsycho.2020.05.010.

42. Li, Y., Long, J., Yu, T., Yu, Z., Wang, C., Zhang, H., and Guan, C. (2010). An EEG-Based BCI System for 2-D Cursor Control by Combining Mu/Beta Rhythm and P300 Potential. IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. *57*, 2495–2505. https://doi.org/10.1109/TBME.2010.2055564.

43. Shi, N., Miao, Y., Huang, C., Li, X., Song, Y., Chen, X., Wang, Y., and Gao, X. (2024). Estimating and approaching the maximum information rate of noninvasive visual brain-computer interface. Neuroimage *289*, 120548. https://doi.org/10.1016/j.neuroimage.2024.120548.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Dataset	Authors	https://data.mendeley.com/datasets/c3mhk3vvh2/1
Software and algorithms		
Python	Python Software Foundation	https://www.python.org/
PsychoPy	Open Science Tools Ltd.	https://psychopy.org/
CURRY 8	NeuroScan Inc.	https://compumedicsneuroscan.com/
Custom code	Authors	https://github.com/Oser-1210/Visual_Tracking_BCI.git
Other		
SynAmps RT	NeuroScan Inc.	https://compumedicsneuroscan.com/
Quik-Cap	NeuroScan Inc.	https://compumedicsneuroscan.com/

RESOURCE AVAILABILITY

Lead contact

Xiaorong Gao (gxr-dea@mail.tsinghua.edu.cn).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Processed data from this study can be obtained at https://data.mendeley.com/datasets/c3mhk3vvh2/1. Raw data is available from the lead contact upon request.
- The code is publicly available at https://github.com/Oser-1210/Visual_Tracking_BCI.git.
- Any additional information required is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Seventeen individuals (6 females, 7 experienced) participated in the study. Each participant was asked to read and sign an informed consent form approved by the Institution Review Board of Tsinghua University before the experiments. All participants received financial compensation at the end of their participation. The participants underwent EEG data collection during the experiment and willingly shared their EEG data without disclosing any personal information. A description of the participants characteristics is provided in the Table S1.

METHOD DETAILS

Stimulus paradigm

Multi-region stimulus

Different from classical speller design, the paradigm aims to track the location of fixation and fulfill cursor control based on neural activity. When fixating on the target location (green circle in Figure 1A), the cursor tracks and automatically moves to the location based on EEG activity. To achieve this, the whole visual field is separated into 8 radial stimulus regions, centering the cursor. The relative position between the stimulus and the cursor remains constant throughout the cursor movement. It should be noted that in this context, "region" specifically pertains to the eight individual radial sectors, while "stimulus" encompasses the overall stimulus. The latter can refer to the entire stimulus paradigm or all stimuli within the field of view.

To achieve continuous control, each region of the stimulus is modulated by a sequence of uniformly-distributed White Noise (WN; see Figure S1). The WN-based modulation belongs to code-modulated visual evoked potential-based BCIs (c-VEP BCIs) and was recently proposed to exhibit multifold advantages over traditional sinusoidal coding.[43] The significant advantages of WN modulation are that the broadband tagging stimulus balances the spectral preference of the primary visual channel, thus avoiding uneven decoding results that prefer a certain class. To put it in the context of continuous control, expressing a preference for a particular radial stimulus could lead to potential directional drift and an increased probability of errors.

Stimulus implementation

In the experiment, the screen was set to 800 × 800 pixels with a black background. On top of the background, the 8 stimulus regions flickered with 1 s of WN sequence and repeated consecutively. The red point, situated at the center of the radial stimulus with a radius of 5 pixels, served as the cursor. The green target was a semi-transparent circle with a radius of 40 pixels. To reduce visual fatigue, the contrast limit of each region was kept to gray (RGB=(127, 127, 127)). This implies that the RGB sequences for each region are obtained by multiplying the WN sequences by (127, 127, 127). To implement the tracking experimental design, the stimulus material (picture form) was generated larger than the screen size (Figure 1B), which means only part of the stimulus is visible on the screen. As each movement proceeded, the stimulus picture and its center were adjusted correspondingly. The entire BCI system, including stimulation, acquisition, and operation, was implemented through PsychoPy.

Experiments materials and design

The stimulus was displayed on a 1920 × 1080 ASUS MG279Q monitor with a 60 Hz refresh rate. EEG data were recorded by Synamps2 system (NeuroScan, Inc.) at a 1000 Hz sampling rate with 62 channels, and 21 of them (Pz, P1/2, P3/4, P5/6, P7/8, POz, PO3/4, PO5/6, PO7/8, Oz, O1/2, and CB1/2) were used for analysis. The recorded EEG data were first down-sampled to 250 Hz, then band-pass filtered to 4–100 Hz, and notch-filtered for the 50 Hz interference.

The experiments contain three parts: training (Stage I, Stage II), testing (fixed tracking task, random tracking task), and application experience (painting, snake game). During the training phase, each participant was instructed to fixate on the green target in both Stage I and Stage II (Figure 2A), without receiving online feedback on cursor movement. The purpose of Stage I is to train a template matching model for calculating correlation coefficient matrix (P; $P \in \mathbb{R}^{1 \times N_r}$, N_r stands for the number of stimulus regions), and P will be employed to predict velocity. In this stage, there were 8 targets, each located on the center line of its respective region, positioned at a distance of 1/3 of the screen width (w_s) from the center red point. The aim of Stage II is to train the corrected velocity weight for generating more efficient velocity for continuous cursor control (see Decoding algorithms for Velocity weight). To achieve this, we introduced 32 targets in this stage (as in Figure 2A), including intersected directions at the boundary between two adjacent regions and a new distance of 1/6 of w_s from the center point (forming inner and outer circle). In both stages, each target was presented 6 times for a duration of 1 s each time.

During the testing phase, while participants fixated on the cued target, the EEG data collected at each step, with a duration of 1 s and no overlap, was used to determine the decoded velocity, which in turn controlled the cursor's movement toward the target. Then, the red cursor moved at the decoded velocity until it reached the target (Figure 2A). The testing phase contains two tasks: fixed tracking task and random tracking task. During the fixed tracking task, the green targets appear at the same 32 positions as in training Stage II. Participants are instructed to fixate on the cued target until the cursor hits the target, during which eye blinking is allowed. If the target is successfully reached within 15 s, the system provides feedback on the time cost and waits for the participant to press space to proceed to the next trial. If not, the system displays a timeout message and moves on to the next trial. At the beginning of each trial, the cursor is refreshed to the center of the screen. The fixed tracking task consists of 3 blocks, with the 32 positions repeated in each block. In the random tracking task, the green target appears at a random position on the screen. At the beginning of each consecutive trial, the cursor starts from the end position of the previous trial. In this task, a total of 12 trials are performed.

During application experience, participants were engaged in using the built-in applications, specifically painting and snake game. In the painting stage, participants were instructed to use their visual gaze as brushstrokes, allowing them to paint at their will. In the gaming stage, we used the downgraded version of continuous decoding, where the screen interface is split into blocks and the cursor moves one grid at a time. The cursor of this game is simplified as the head of the snake, and target is shown as the food. Once the snake obtains the food, its body length increases by the length of one block. Participants were instructed to increase the body length while avoiding collisions with the screen edges or the snake's own body. Following the game, participants were awarded a score that corresponded to the snake's length. The decoding algorithm of the snake game employs a dynamic stopping strategy based on confidence estimation (see Decoding algorithms for Confidence estimation based on hypothesis testing).

Decoding algorithms

Spatial filtering

We chose Task-Related Component Analysis (TRCA) as the template matching method.[24] The goal of TRCA is to find a set of spatial filters that maximize the task-related components by optimizing the inter-trial consistency. The optimization goal can be translated as the maximization of inter-trial covariance after applying the spatial filters, expressed as:

$$w^T S w = \sum_{\substack{h_1, h_2 = 1 \\ h_1 \neq h_2}}^{N_t} \sum_{j_1, j_2 = 1}^{N_c} w_{j_1} w_{j_2} \text{Cov}\left(x_{j_1}^{(h_1)}(t), x_{j_2}^{(h_2)}(t)\right)$$

where h means different trials and j means different channels. N_t is the number of trials, and N_c is the number of channels. The optimization problem is solved under the constraint of:

$$w^T Q w = \sum_{j_1, j_2 = 1}^{N_c} w_{j_1} w_{j_2} \mathrm{Cov}\left(x_{j_1}(t), x_{j_2}(t)\right)$$

the final spatial filter can be determined as follows:

$$\widehat{w} = \underset{w}{\mathrm{argmax}} \frac{w^T S w}{w^T Q w}$$

In the typical classification task, template matching is made by comparing the correlation coefficient matrix between spatial weighted evoked template and single trial response (Figure 2B). And the final prediction is made by selecting the stimulation class corresponding to the maximum correlation coefficient (ρ_{max}). It's important to note that, in the TRCA method, due to the use of filter banks, the final value of each correlation coefficient (ρ) is obtained by weighting the correlation coefficients (using r as a distinct reference) between the sub-filtered response and the template. Consequently, the values of ρ s do not conform to the standard range of -1 to 1.

$$\rho = \sum_{m = 1}^{N_m} a(m) \cdot r_m$$

where m means different sub-bands and N_m is the total number of sub-bands. $a(m)$ is defined as $a(m) = m^{-1.25} + 0.25$.

Velocity prediction

The key distinction from discrete classification to continuous tracking is to change template matching based on ρ_{max} to coefficient projection based on the entire coefficients (Figure 2B). To produce continuous movement, we map the correlation coefficients (ρ s) of distinct patterns to ρ vectors starting from the cursor center and point to their corresponding 8 directions. The sum of all ρ vectors, as we call projection result in this study, represents the intended direction of cursor movement. The change from ρ_{max} to ρ s promises a significant improvement of decoding algorithm. By leveraging the correlation coefficients from other stimulus regions, we quantified the combined effects of various stimulus regions in the visual field. It's worth noting that the correlation coefficients can take on negative values, and the aforementioned projection method conducts unidirectional projection. Consequently, ReLU is employed to reset negative values to zero. However, the projection result doesn't align with practice scenarios, and due to variations in stimulus responses among participants, it lacks universality and may cause direction drift. To minimize the projection error and thrive for ideal tracking performance, we introduced the velocity weight for correction.

Velocity weight

The velocity weight serves two main purposes. One is to ensure that the projection result corresponds to a real velocity that matches the screen size, while the other is to individually adjust and correct the magnitude and direction of the decoded velocity for each participant. The velocity vector (v) can be calculated as the dot product of the correlation coefficient matrix (P) and velocity weight (V_w; $V_w \in \mathbb{R}^{N_r \times 2}$, 2 represents the Cartesian coordinates x and y).

$$v^T = P \cdot V_w$$

We can align the decoded velocity with the screen size by using an initial velocity weight (V_w^{init}) distributed in a circular pattern with a radius of $1/3$ of the screen width (w_s) divided by 2 (the average length of projection result).

$$V_w^{init} = \begin{pmatrix} 1 & 0 \\ \sqrt{2}/2 & \sqrt{2}/2 \\ 0 & 1 \\ -\sqrt{2}/2 & \sqrt{2}/2 \\ -1 & 0 \\ -\sqrt{2}/2 & -\sqrt{2}/2 \\ 0 & -1 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{pmatrix} \cdot \frac{w_s}{6}$$

To facilitate a more natural effect, during the practical experiment, the velocity gradually diminishes to zero within a time step (1 s). The average velocity of this process corresponds to the decoded velocity.

Furthermore, to minimize prediction error between intended velocity and decoded velocity (Figure 2A) and solve the issue of subject differences, we trained a corrected velocity weight (V_w^*) by least square method for each participant, using data from training Stage II:

$$V_w^* = \left(D^T D\right)^{-1} D^T I$$

where D is the decoded correlation coefficient matrix for all trials ($D \in \mathbb{R}^{N_t \times N_r}$, N_t is the number of trials and N_r is the number of stimulus regions), and I is the intended velocity matrix for all trials ($I \in \mathbb{R}^{N_t \times 2}$), where intended velocity is calculated as the distance from the target to the center red point divided by one time step (1 s). If there is no special prompt, the velocity weight in Results refers to the corrected (or trained) one.

Confidence estimation based on hypothesis testing

In the snake game, where the snake advances one grid at a time using classification methods, we have integrated a dynamic stopping strategy. This strategy helps determine whether the snake should move or not, aiming to minimize the likelihood of game failures caused by incorrect judgments. The strategy evaluates prediction confidence based on the hypothesis testing framework that has been used in the previous studies. Based on the correlation coefficient matrix above, we calculate the confidence level between the maximum correlation coefficient and the remaining correlation coefficients. We can decide whether to reject the hypothesis and, accordingly, determine whether the snake should move by setting a threshold for the p value.

BCI performance evaluation

Different from typical BCI speller which use ITR for evaluation metrics, we employ Fitt's ITR to properly evaluate BCI performance in the continuous task.[33] The Fitt's ITR is related to three parameters, including the target size, distance to the target, and the time cost.

$$\text{Fitt's ITR} = \frac{\log_2 ((D+S)/S)}{T}$$

where S represents the target size, D represents the distance to the target, and T represents the time cost (see Figure S2). The success rate is computed as the percentage of trials in which the cursor reaches the target within the time limit. The time to target is the time required for the cursor to move from the center or start position to the target. The unsuccessful trials are not involved in the calculation of Fitt's ITR and time to target, while they are presented as success rate.

Other than Fitt's ITR, we defined two errors as the metrics of the velocity prediction (Figure 3C). One is angular error, which measures the angular difference between the decoded velocity and the intended velocity. The other is vector error, which is calculated using the length of the difference between the intended velocity vector and the decoded velocity vector to divide the length of the intended velocity vector. The error calculation is based on the decoded velocity from the first step of each trial in the fixed tracking task. By combining Fitt's ITR and velocity errors, we can yield objective measurements for the decoding algorithm and the resulting tracking performance.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analysis were conducted using Python 3.9.13. The packages involved include scipy 1.8.0 and statsmodels 0.13.2. Comprehensive details of the statistical analyses have been provided in the main text, figure captions, and methods. These details encompass the statistical tests employed, the number of samples, the mean and standard deviation of the statistics, and the p value. The significance test methods include t test, two-way ANOVA, with the correlation coefficient typically being the Pearson correlation coefficient. The use of the Spearman correlation coefficient has been elucidated in the main text.