scientific reports

OPEN

The effect of different depth planes during a manual tracking task in three-dimensional virtual reality space

Hyeonseok Kim[1], Yasuharu Koike [2], Woong Choi[3]✉ & Jongho Lee[4]✉

Unlike ballistic arm movements such as reaching, the contribution of depth information to the performance of manual tracking movements is unclear. Thus, to understand how the brain handles information, we investigated how a required movement along the depth axis would affect behavioral tracking performance, postulating that it would be affected by the amount of depth movement. We designed a visually guided planar tracking task that requires movement on three planes with different depths: a fronto-parallel plane called ROT (0), a sagittal plane called ROT (90), and a plane rotated by 45° with respect to the sagittal plane called ROT (45). Fifteen participants performed a circular manual tracking task under binocular and monocular visions in a three-dimensional (3D) virtual reality space. As a result, under binocular vision, ROT (90), which required the largest depth movement among the tasks, showed the greatest error in 3D. Similarly, the errors (deviation from the target path) on the depth axis revealed significant differences among the tasks. Under monocular vision, significant differences in errors were observed only on the lateral axis. Moreover, we observed that the errors in the lateral and depth axes were proportional to the required movement on these axes under binocular vision and confirmed that the required depth movement under binocular vision determined depth error independent of the other axes. This finding implies that the brain may independently process binocular vision information on each axis. Meanwhile, the required depth movement under monocular vision was independent of performance along the depth axis, indicating an intractable behavior. Our findings highlight the importance of handling depth movement, especially when a virtual reality situation, involving tracking tasks, is generated.

From a motor control viewpoint, manual tracking movements involve both feedback and feedforward control[1]. The extent to which the brain relies on feedforward control for successful performance depends on several factors, including age[2], disorders[3], target acceleration[4], and target speed[5], involving sub-movements or intermittency[6–8]. Unlike reaching movements, which are likely carried out by initial feedforward control transitioning into feedback control when closer to a target, visually guided tracking tasks require reliance on perpetual processing of spatial information involving visual feedback. Additionally, many manual tracking tasks have been performed in two-dimensional space[9–13]. Those performed in three-dimensional (3D) space have not focused on the effect of depth, as the required trajectory has been somewhat sophisticated[14]. However, depth information of the target should be investigated to understand how the brain plans and generates motor commands to perform tracking movements, as depth perception requires intricate computation in the brain[15]. In other words, the extent to which depth information contributes to performing a manual tracking movement in 3D space remains to be elucidated.

Depth perception is associated with considerable information processing in the brain. Under binocular vision, binocular disparity has been regarded as an important factor in depth perception[16]; therefore, several models have been suggested to explain how the brain can exploit this information[17]. Binocular vision is also related to binocular fusion, which can influence depth perception[18]. Such information is only available under binocular vision and could have advantages in motor control[19]. Relevant brain activity has also been reported, including

[1]Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California San Diego, La Jolla, CA 92093, USA. [2]Institute of Innovative Research, Tokyo Institute of Technology, Yokohama 226-8503, Japan. [3]College of ICT Construction & Welfare Convergence, Kangnam University, Yongin 16979, Republic of Korea. [4]Department of Clinical Engineering, Komatsu University, Komatsu 923-0961, Japan. ✉email: wchoi@kangnam.ac.kr; jongho.lee@komatsu-u.ac.jp

1

activity in the dorsal pathway for depth perception by fusing relative motion and binocular disparity[20], neuronal activity associated with depth based on motion parallax[21], and brain activity in the posterior parietal cortex and dorsal stream associated with the kinetic depth effect[22]. Models explaining depth perception under monocular vision have also been suggested[23]. Since depth perception depends on various types of available information[24], it is important to investigate how this information acts at the performance level.

Thus, we focused on the effect of depth information in a visually guided tracking task, which could be useful for learning complex motor skills, especially in a 3D virtual reality (VR) space. 3D interactions have been conducted in situations implemented by different systems, including infrared imaging[25], augmented reality[26], and tablet personal computers[27]. We adopted a VR environment to create a situation with a target that spatially corresponded to real-world coordinates for a realistic experience[28]. In VR space, a visual target is completely under our control, permitting the tracer to overlap the target, unlike in the real world, where construction of the environment necessitates a physical entity circularly moving at a constant speed.

If depth information influences tracking performance in a complicated manner, it should be considered when designing training programs for facilitating motor skills through visually guided tracking tasks in 3D VR space. Although a significant difference in tracking performance between the fronto-parallel plane and sagittal plane was reported in terms of an error on the lateral axis[28], the contribution of depth information to the performance of the task was not explained.

This study aimed to investigate how required movement along the depth axis would influence 3D tracking performance, postulating that the amount of depth movement would affect performance. We designed a visually guided planar tracking task that requires movement on three planes with different depth spaces: a fronto-parallel plane, a sagittal plane, and a plane rotated by 45° with respect to the sagittal plane. In particular, tasks on the fronto-parallel and sagittal planes require movement on two independent axes in the eye-centered coordinate system. However, the plane rotated by 45° with respect to the sagittal plane requires the brain to process three-dimensional information. By adopting this plane, we could determine whether depth information interferes with information processing on the other axes in behavioral performance, regardless of the fundamental processing of the brain. In order to understand the comprehensive effect depth information has on 3D tracking performance, we experimented under two vision environments (binocular and monocular visions) because this category is involved in processing depth information within the brain and can impact tracking performance.

Methods

Participants

Fifteen male participants were recruited for this experiment. Table 1 presents the demographic information of the participants. Their ages' mean and standard deviation (SD) were 20.1 and 0.6, respectively. Two participants were left-handed, and the others were right-handed. All participants provided written informed consent before the experiment. This study was approved by the Institutional Review Board of National Institute of Technology, Gunma College, and conducted in accordance with the Declaration of Helsinki.

Experimental apparatus

In the experiment, participants were asked to perform a visually guided tracking task[28] with HTC VIVE (HTC Corporation, Taipei, Taiwan) in an immersive 3D VR space, which was implemented using Unity software (Unity Technologies, San Francisco, CA, USA). During the experiment, the participants wore a VIVE head-mounted display (HMD) capable of a refresh rate of 90 Hz, which was used for objects in the VR space. The HMD offered a 110-degree field of view with a 2880×1600 pixels resolution. The HTC VIVE system adopts the Lighthouse tracking system[29] with a position precision of 2 mm. The position of the hand holding the controller and its

Participant	Age	Sex	Handedness
1	20	M	Left
2	20	M	Right
3	20	M	Right
4	20	M	Right
5	20	M	Right
6	19	M	Right
7	21	M	Right
8	21	M	Right
9	19	M	Right
10	21	M	Right
11	20	M	Right
12	20	M	Right
13	20	M	Right
14	21	M	Right
15	20	M	Left

Table 1. Demographic information of the participants.

inclination were visualized with a pole parallel to the controller that had a sphere at one end to indicate the hand position in the 3D VR space, as shown in Fig. 1, so that the participants could keep track of a target moving at a constant speed by controlling the tracer. A red sphere with a diameter of 3 cm was presented as the target, and a yellow sphere with a diameter of 2 cm was presented at the end of the tracer. The length of the pole was 20 cm, but its color was white, unlike that in Fig. 1. The diameter of the invisible target trajectory was 30 cm. In this study, the X-, Y-, and Z-axes corresponded to the lateral, vertical, and depth dimensions.

Experimental procedure

In this study, we designed an experiment to investigate the motor control mechanism during a tracking task in 3D VR space. As shown in Fig. 2, the experiment involved three kinds of tasks with different depth planes: ROT (0), in which the target was presented by moving in a circle on a fronto-parallel plane; ROT (45), in which the target moved in a circle on a plane rotated by 45° with respect to the sagittal plane; and ROT (90), in which the target moved in a circle on the sagittal plane. In other words, the ROT (0) task required lateral and vertical movements, whereas the ROT (90) task required vertical and depth movements. In contrast, ROT (45) required concurrent lateral, vertical, and depth movements. In addition, this experiment comprised two vision conditions: binocular vision condition, where participants performed the tasks without having their vision blocked in any way, and monocular vision condition, where participants performed the tasks with one side of their vision obstructed.

Before the experiment, the participant sat on a chair and wore the HMD, followed by a calibration session in which the individualized initial position of the target was varied across the height of each participant. The arm length was set to minimize variations due to the diversity of physical characteristics[28]. The participants
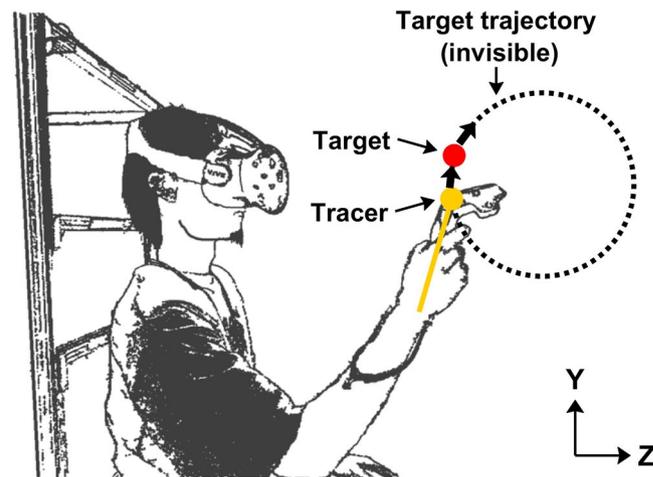
Figure 1. Experimental setup. Participants sat and manipulated a VIVE controller with their dominant hand to move the tracer during the experiment. A red sphere with a diameter of 3 cm was presented as a target, and a yellow sphere with a diameter of 2 cm was presented for the end of the tracer in the VR space. The length of the pole was 20 cm. The actual presented color of the pole in the VR space was white. The target trajectory was invisible in the actual experiment. The diameter of the target trajectory was 30 cm. The X, Y, and Z axes correspond to the lateral, vertical, and depth dimensions. This figure illustrates an image of the YZ plane, that is, the sagittal plane.
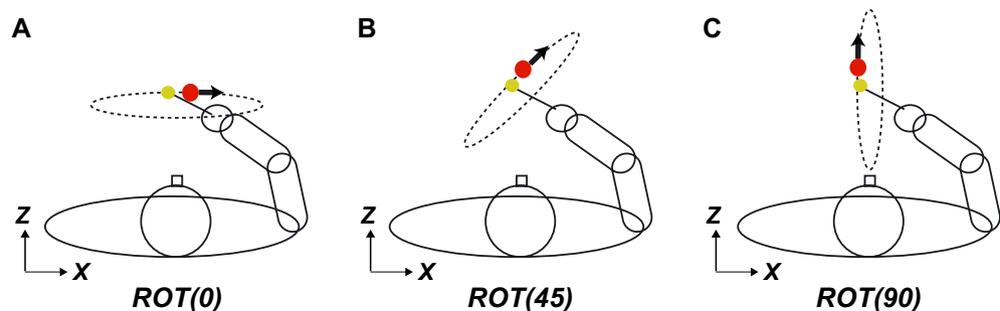
Figure 2. Experimental task. The tracking task in this experiment comprised three kinds of tasks: ROT (0), ROT (45), and ROT (90). For the ROT (0) task, the target moved in a circle on a fronto-parallel plane. For the ROT (45) task, the target moved in a circle on the plane rotated by 45° with respect to the sagittal plane. For the ROT (90) task, the target moved in a circle on the sagittal plane. Participants were instructed to track the target moving at a constant speed of 0.5 Hz.

manipulated the controller with their dominant hand in all tasks. For the monocular condition, participants wore an eye patch to block one side of their eye, identical to the opposite side of their dominant hand, prior to the experiment. For each trial, a sound was generated three times for 3 s, with an interval of 1 s. After 3 s, the target began to move at a constant tangential velocity (0.5 Hz at which the target moved in a circle in 2 s), following the invisible circle in a clockwise direction, as shown in Figs. 1 and 2. The target was stopped after moving in a circle three times. The participants were instructed to match the center of the tracer to that of the target during the task. There were three trials for each task after each trial, which was discarded from analysis. Each participant conducted 12 trials, including a practice trial for each vision condition.

Data analysis

During the movement task, we recorded the positions of the target and tracer in 3D space at a 90 Hz sampling rate. To evaluate performance, we calculated the position error in 3D space and the absolute value of the error on each axis, which are defined as follows:

$$error\ in\ 3D\ [mm] = \sqrt{(T_x - t_x)^2 + (T_y - t_y)^2 + (T_z - t_z)^2} \tag{1}$$

$$error_{X-axis}[mm] = |T_x - t_x| \tag{2}$$

$$error_{Y-axis}[mm] = |T_y - t_y| \tag{3}$$

$$error_{Z-axis}[mm] = |T_z - t_z| \tag{4}$$

where T_x, T_y, and T_z represent the X, Y, and Z coordinates of the target position, respectively, and t_x, t_y, and t_z represent the X, Y, and Z coordinates of the tracer's position, respectively. We averaged these values over the performance time for the analysis.

For each condition and parameter (the abovementioned errors), we performed a one-way repeated measures analysis of variance (ANOVA) (three levels for the plane factor: *ROT (0)*, *ROT (45)*, and *ROT (90)*) using SPSS Statistics (IBM, Armonk, NY, USA). The sphericity assumption was verified using Mauchly's test. For Mauchly's test, $p < 0.05$, the p-value for the ANOVA was corrected using Greenhouse–Geisser. A post-hoc test was conducted through pairwise comparisons using Bonferroni correction. We considered comparisons yielding $p < 0.05$ to be statistically significant and comparisons yielding $p < 0.005$ to be highly statistically significant.

Ethics declarations

All participants provided written informed consent before the experiment. This study was approved by the Institutional Review Board of National Institute of Technology, Gunma College, and conducted in accordance with the Declaration of Helsinki.

Results

Tracking performance in the binocular vision condition

Figure 3 shows an example of tracking movement in the binocular vision condition. Although certain errors in the depth axis were observed in the ROT (45) and ROT (90) tasks, the tracer generally tracked the target with small errors. Figure 4 shows the 3D errors for all types of tasks in the binocular vision condition. The mean and SD of the errors in 3D were 22.9 ± 4.3, 22.2 ± 3.9, and 26.2 ± 4.4 for ROT (0), ROT (45), and ROT (90), respectively. In the comparison between the ROT (0) and ROT (45) tasks, no significant difference in tracking performance was observed (corrected $p = 1$). However, we observed a significantly large error in 3D in the ROT (90) task (corrected $p = 0.009$ for the comparison between the ROT (0) and ROT (90) tasks; corrected $p = 0.001$ for the ROT (45) and ROT (90) tasks).

As we found significant differences in the 3D errors, we investigated the errors on each axis. Figure 5 shows the errors on each axis for the three types of tasks in the binocular vision condition. The errors on the X-axis were 12.3 ± 2.3, 9.15 ± 1.9, and 4.52 ± 1.0 for the ROT (0), ROT (45), and ROT (90) tasks, respectively. These differences were significant for all comparisons (corrected $p < 0.001$ for the ROT (0) and ROT (45) tasks; corrected $p < 0.001$ for the ROT (45) and ROT (90) tasks; and corrected $p < 0.001$ for the ROT (0) and ROT (90) tasks). For the Y-axis, the errors were 11.0 ± 1.6, 10.7 ± 1.9, and 11.9 ± 2.0 for the ROT (0), ROT (45), and ROT (90) tasks, respectively. The error in the ROT (90) task was the largest, but not significant, as tested with ANOVA ($p = 0.056$ corrected using Greenhouse–Geisser). For the Z-axis, the errors were 11.1 ± 3.8, 13.8 ± 3.2, and 20.1 ± 4.5 for the ROT (0), ROT (45), and ROT (90) tasks, respectively. Among the tasks, performance in the ROT (90) task was the worst, while performance in the ROT (0) task was the best. The differences were statistically significant (corrected $p = 0.023$ for the ROT (0) and ROT (45) tasks; corrected $p < 0.001$ for the ROT (45) and ROT (90) tasks; and corrected $p < 0.001$ for the ROT (0) and ROT (90) tasks).

Tracking performance in the monocular vision condition

Figure 6 shows an example of tracking movement in the monocular vision condition, indicating that the tracer on the depth axis in the ROT (90) task frequently deviates from the correct path. Figure 7 shows the 3D errors for all types of tasks in the monocular vision condition. The mean and SD of the errors in 3D were 59.6 ± 34.5, 59.3 ± 32.9, and 56.4 ± 21.1 for the ROT (0), ROT (45), and ROT (90) tasks, respectively. The differences in 3D errors for all types of tasks were not statistically significant, as tested with ANOVA ($p = 0.545$).
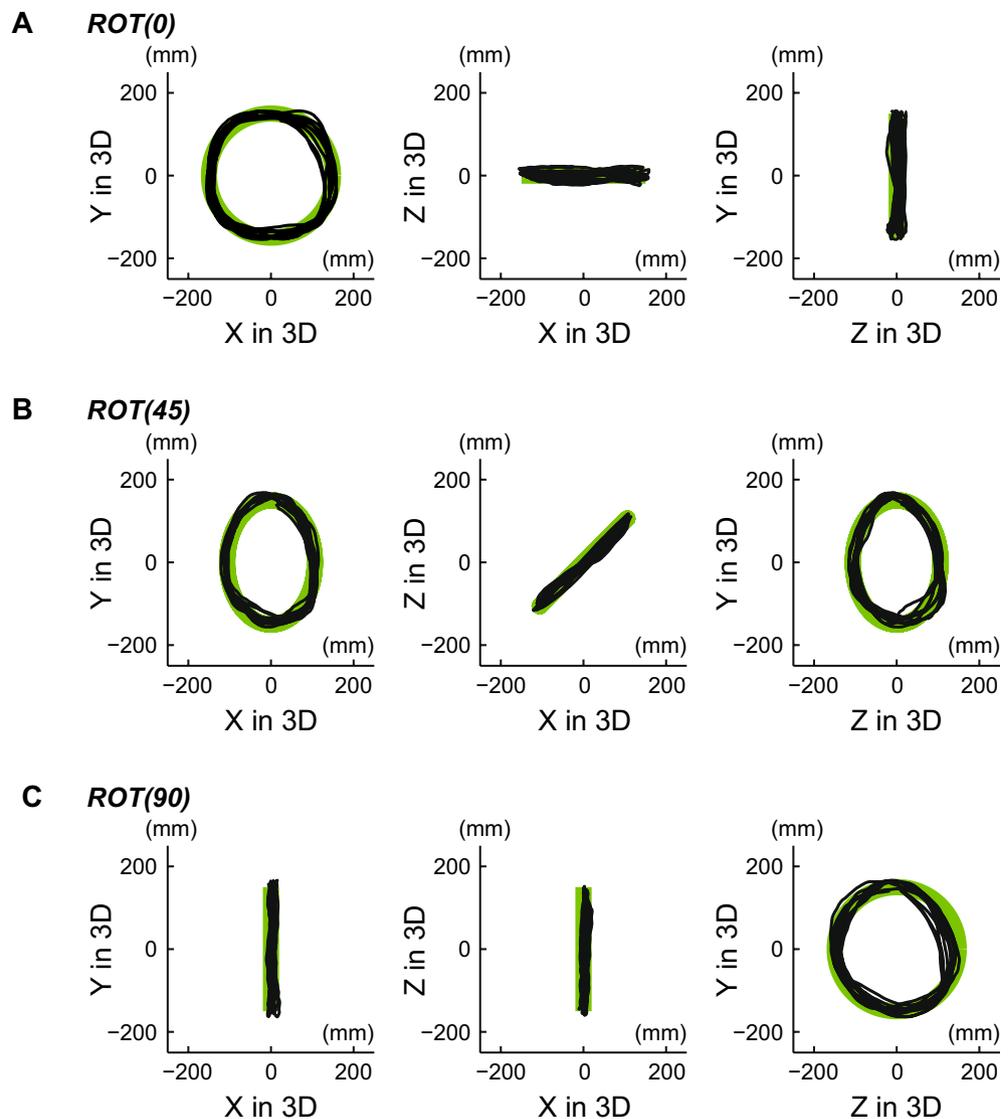
Figure 3. An example of tacking movement in the binocular vision condition. Each row represents the ROT (0), ROT (45), and ROT (90) tasks, respectively. The green line represents the movement of the target moving at a constant speed, and the black line represents the trajectory of the tracer.

Figure 8 shows the errors on each axis for the three types of tasks in the monocular vision condition. The errors on the X-axis were 21.0 ± 7.1, 20.9 ± 6.8, and 10.8 ± 3.9 for the ROT (0), ROT (45), and ROT (90) tasks, respectively. The performance in the ROT (90) task was significantly better than that in the other tasks (corrected $p < 0.001$ for the ROT (0) and ROT (90) tasks; corrected $p < 0.001$ for the ROT (45) and ROT (90) tasks). However, the difference between the ROT (0) and ROT (45) tasks was not significant (corrected $p = 1$ for the ROT (0) and ROT (45) tasks, respectively). For the error on the Y-axis, the performance in all tasks was the same as tested with ANOVA ($p = 0.48$). The mean and SD of the errors on the Y-axis were 23.5 ± 9.3, 24.8 ± 10.7, and 25.9 ± 9.1 for the ROT (0), ROT (45), and ROT (90) tasks, respectively. Similar to the Y-axis, the errors on the Z-axis were the same in all tasks, as tested with ANOVA ($p = 0.822$). The errors on the Z-axis were 59.6 ± 34.5, 59.3 ± 32.9, and 56.4 ± 21.1 for the ROT (0), ROT (45), and ROT (90) tasks, respectively.

Discussion

In this study, we quantitatively evaluated the tracking performance of circular movements using three types of tasks with different depth planes in 3D VR space. In particular, we adopted two vision conditions (binocular and monocular vision) and investigated the effect of the required movement along the depth axis during tracking tasks in 3D VR space. Under binocular vision, the ROT (90) task, which requires the largest depth movement, showed the greatest error in 3D among all tasks (Fig. 4). Likewise, the errors on the depth axis revealed significant differences among the tasks (Fig. 5). Under monocular vision, significant differences in errors were observed only on the x-axis (Fig. 8). These findings indicate that the amount of required depth movement under binocular vision determines the depth error independently of the other axes, implying that the brain may independently
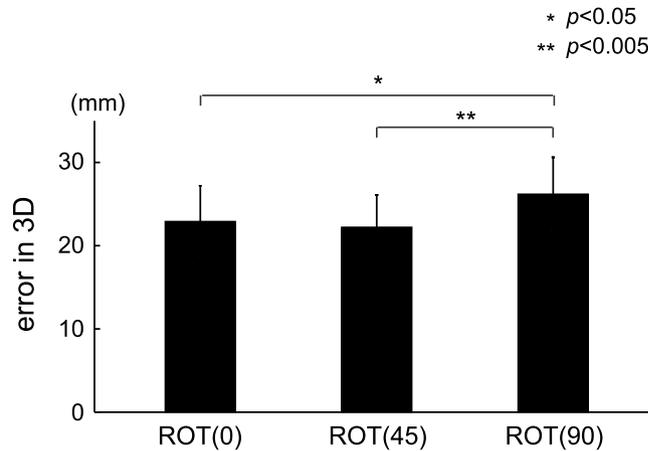
Figure 4. Errors in 3D for the three tasks in the binocular vision condition. We found significant differences in errors between the ROT (0) and ROT (90) tasks and between the ROT (45) and ROT (90) tasks (t(14) = 3.608, corrected p = 0.009 for the comparison between the ROT (0) and ROT (90) tasks; t(14) = 4.619, corrected p = 0.001 for the ROT (45) and ROT (90) tasks). The difference in error in 3D between the ROT (0) and ROT (45) tasks was not significant (t(14) = 0.987, corrected p = 1).

process binocular vision information on each axis. Meanwhile, the required depth movement under monocular vision was independent of performance along the depth axis, indicating an intractable behavior.

For the binocular vision condition, the error in 3D in the ROT (90) task was significantly high, suggesting that a task requiring movement on the depth axis is demanding. When we looked at the decomposed errors along each axis, the ROT (0) task involved the largest error on the X-axis, whereas the ROT (90) task involved the largest error on the Z-axis. Errors in the ROT (45) task were between those in the ROT (0) and ROT (90) tasks, regardless of the axis, implying that information on each axis may be independently processed in the brain. In a previous circular tracking task, a significantly different error on the X-axis between the ROT (0) and ROT (90) tasks was reported[28], in line with our current findings. The errors on the Y-axis were the same, regardless of the rotation. Given that the distance traversed on the Y-axis is the same on all tasks, Y-axis performance is not influenced by information regarding the other axes, supporting the idea that the brain may process information regarding this axis independently (Supplementary Tables).

For the monocular vision condition, errors on the X-axis were involved in the significant difference between the ROT (90) task and either the ROT (0) or ROT (45) task. In addition, the error was not significantly different between the ROT (0) and ROT (45) tasks, suggesting that the amount of horizontal information required to perform the task was not proportional to actual performance. On the depth axis, although no significant differences were observed, the mean error was approximately 60 mm, which is three times greater than the errors in the other axes; this finding was supported by those of a previous study reporting that people tend to underestimate the distance of objects under monocular vision[30,31]. The insignificant differences were attributed to this high variance. Thus, information processing within the brain under monocular vision is more complicated than that under binocular vision.

This difference between monocular and binocular vision may result from several types of information available only under binocular vision, which might include binocular vision involving a better understanding of the properties of a target object[32]. The advantage of binocular vision is associated with spatial accuracy in completing tasks[33]. Furthermore, its superiority has been reported in several motor tasks involving catching[34] or walking[35] as well as when performing tasks that require complex motor skills, such as a handspring[36] or handling tools[37], which supports the validation of our results. In addition, we observed a floor effect of monocular vision that caused large errors in the depth axis across all groups, with insignificant differences when performing a tracking task. In a previous study, the floor effect of monocular vision on tracking performance was observed by blurred vision[38]. These might be used for quantifying the other kinds of capabilities of vision that suffer by depriving one side of vision, considering that performance could be maintained by substituting information only available under binocular vision with that available under monocular vision[39]. Our study revealed that binocular vision outperformed monocular vision in basic tracking tasks.

This study observed only behavioral performance. However, several neural activities correlate with depth processing, including V1 for depth processing[40], V4 for binocular disparity[41], and V2 for processing relative disparity[42]. Moreover, out of several kinds of stimuli, the disparity stimulus has been reported to induce the strongest response in the human visual cortex[43], supporting that positional information is extraordinarily important for visual processing and was used in our study. Moreover, when the primary visual cortex was optogenetically suppressed, jump performance deteriorated in both binocular and monocular vision, which is directly linked to the attribution of behavioral performance to neural activity[39]. In addition to the visual cortex, the corpus callosum is associated with midline depth perception[44]. Studies involving non-human primates have also demonstrated the relationship between depth processing and neural response. For instance, the macaque parietal cortex, related to depth processing, was mainly activated during movement execution rather than during preparation[45]. Additionally, movement direction in a planar reaching task, where movement direction determined
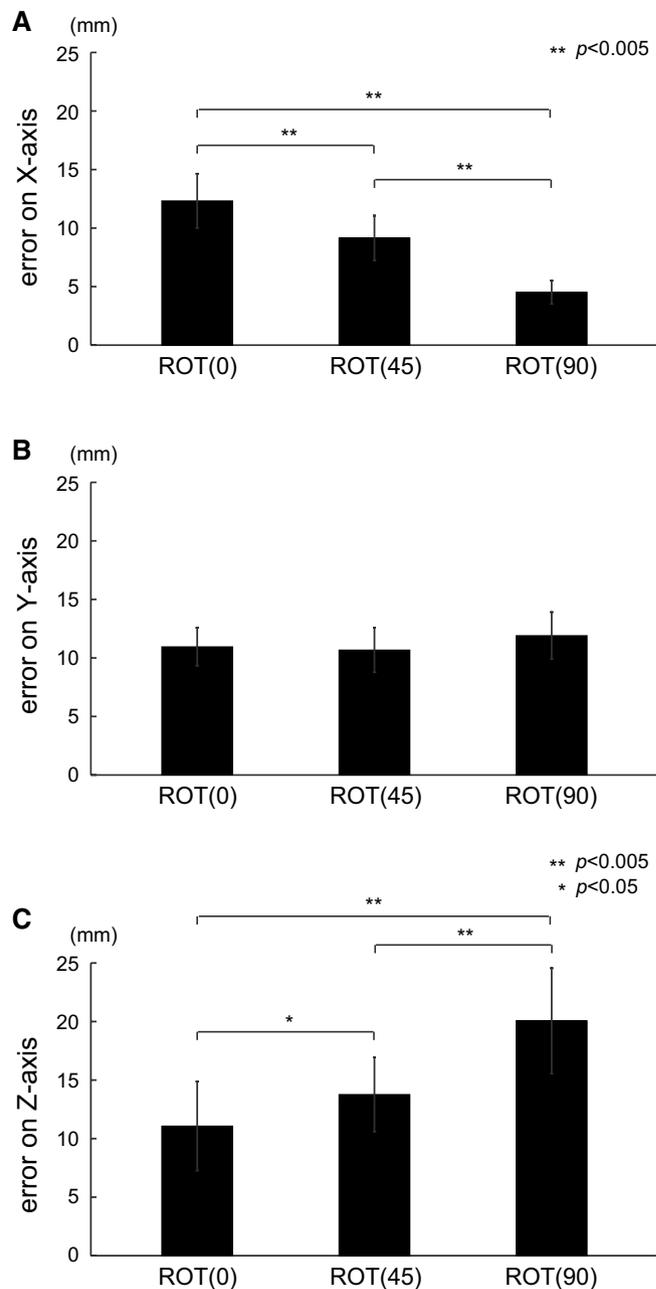
Figure 5. Errors on each axis for the three tasks in the binocular vision condition. Each row represents the X-, Y-, and Z-axis errors, respectively. The differences in errors on the X-axis were significant (t(14) = 7.798, corrected $p < 0.001$ for the ROT (0) and ROT (45) tasks; t(14) = 8.235, corrected $p < 0.001$ for the ROT (45) and ROT (90) tasks; t(14) = 12.536, corrected $p < 0.001$ for the ROT (0) and ROT (90) tasks). The differences in errors on the Y-axis were not significant (F(1.283, 17.958) = 3.887, $p = 0.056$ corrected by Greenhouse–Geisser). For the Z-axis, all the differences in errors were statistically significant (t(14) = 3.117, corrected $p = 0.023$ for the ROT (0) and ROT (45) tasks; t(14) = 6.553, corrected $p < 0.001$ for the ROT (45) and ROT (90) tasks; t(14) = 8.283, corrected $p < 0.001$ for the ROT (0) and ROT (90) tasks).

differing amounts of required movements along the depth axis, was involved in dorsal premotor cell activity[46], and V6A neurons[47]. Direction on a fronto-parallel plane was also associated with neuronal activities in the premotor or primary motor cortex[48], inferring that direction is independent of depth processing. Investigations in the PE area in macaques revealed segregated processing of depth and direction even though the area was involved in the processing of both kinds of information[49]. These studies support the idea that our results can be attributed to neural activity.

Several factors should be considered in future studies. Because visual information, including depth information, is most likely to be used more during feedback control than during feedforward control, how depth information is used during tracking movement should be considered. Since tracking performance is mainly

A *ROT(0)*
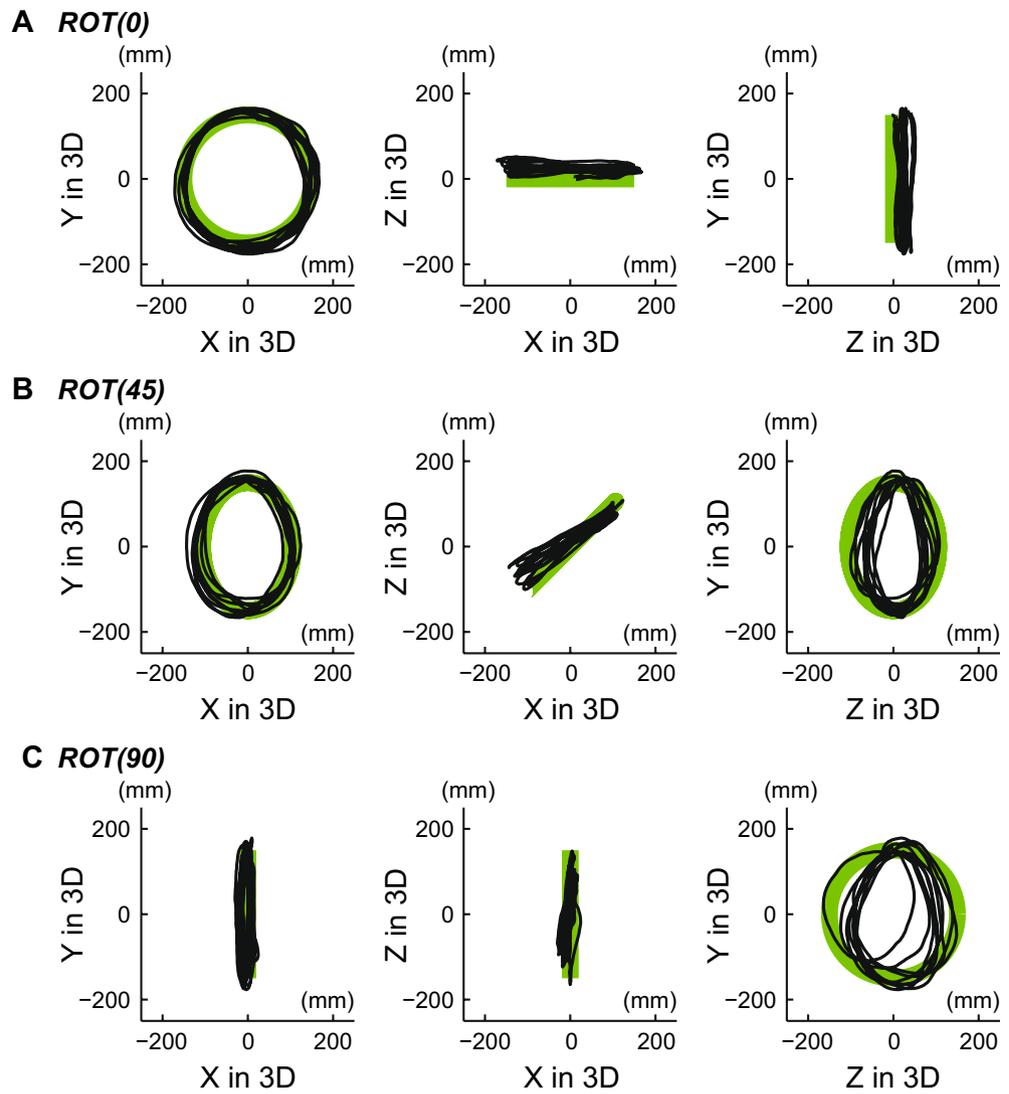
B *ROT(45)*

C *ROT(90)*

Figure 6. An example of tracking movement in the monocular vision condition. Each row represents the ROT (0), ROT (45), and ROT (90) tasks, respectively. The green line represents the target's movement at a constant speed, and the black line represents the trajectory of the tracer. The tracer along the depth axis in the ROT (90) task frequently deviated from the correct path.
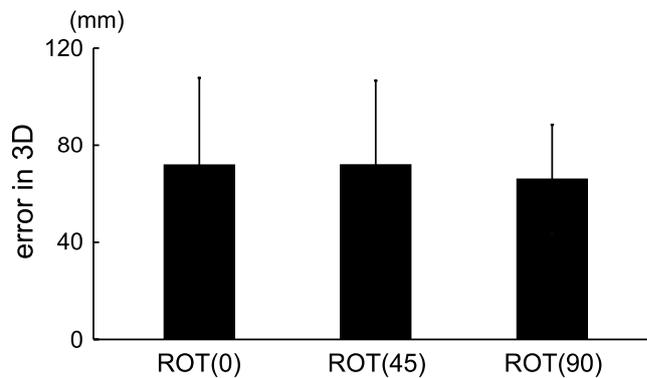
Figure 7. Errors in 3D for the three tasks in the monocular vision condition. None of the differences in errors in 3D were statistically significant ($F(2, 28) = 0.621, p = 0.545$).
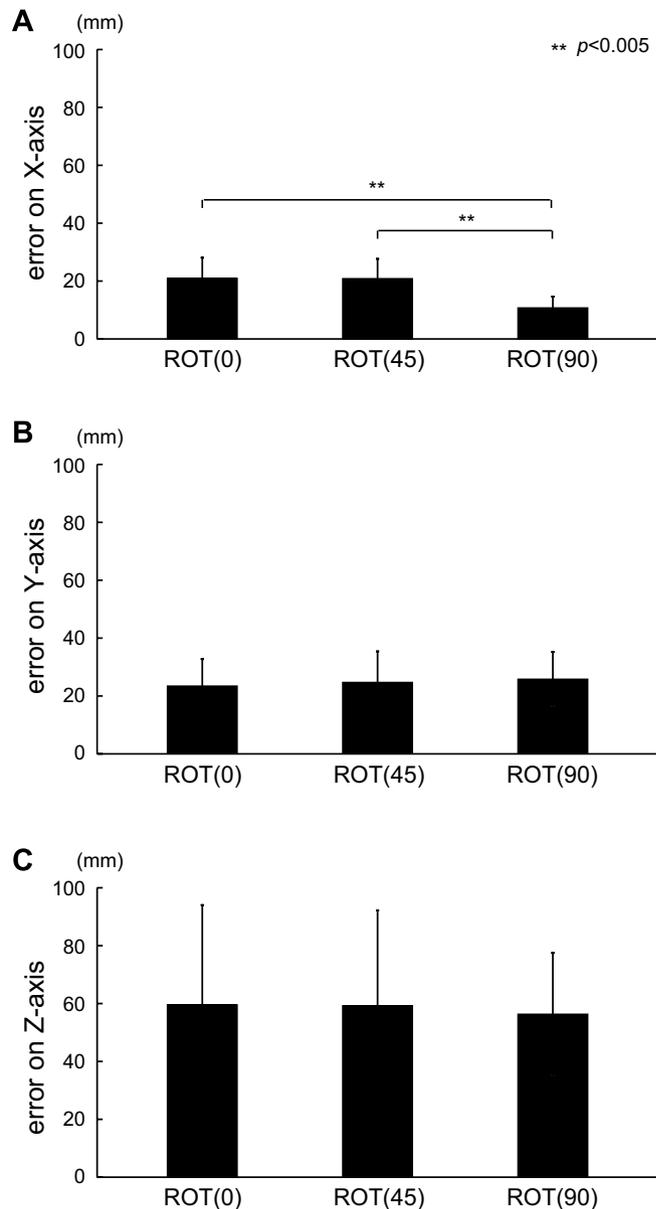
Figure 8. Errors on each axis for the three tasks in the monocular vision condition. Each row represents the X-, Y-, and Z-axis errors, respectively. The error on the X-axis in the ROT (90) task was significantly greater than those in the other tasks (t(14) = 4.554, corrected p < 0.001 for the ROT (0) and ROT (90) tasks; t(14) = 5.507, corrected p < 0.001 for the ROT (45) and ROT (90) tasks). However, the difference between the ROT (0) and ROT (45) tasks was not significant (t(14) = 0.039, corrected p = 1 for the ROT (0) and ROT (45) tasks). For the errors on the Y-axis, a significant difference among tasks was not observed (F(2, 28) = 0.755, p = 0.48). Similar to the Y-axis, the errors on the Z-axis were the same in all tasks (F(2, 28) = 0.197, p = 0.822).

governed by feedforward control, and feedback control is sometimes used for error correction[50], the feedback acts' frequency, duration, and phase should be investigated. In addition, we adopted a constant target speed, but this speed affects motor control in a tracking task[51,52]. Humans' general binocular movement is adjusted to the natural environment[53]; therefore, eye movement during the task could also be measured to see how different the required eye movement for a task is from the adjusted eye movement in a future study. It has been reported that the predictability of a target could modify the control strategy in tracking tasks[1,54]. Although these factors make it difficult to investigate tracking movement, as a first step, we found that the brain might process the information on each axis independently under binocular vision. Moreover, the process of error generation may involve more aspects, and may be possible to investigate from biomechanical and physiological standpoints.

In this study, we investigated how behavioral tracking performance would be affected by the required depth information so that we could understand how the brain deals with information on each axis. We observed that the errors in the lateral and depth axes were proportional to the required movement on the lateral and depth axes

under binocular vision. The tracking performance was unpredictable based on the required lateral movement under monocular vision. Despite its intricate information-processing mechanisms, we confirmed that depth information under binocular vision was independently processed from that in the other axes, whereas monocular vision involved intractable behavior. Therefore, tracking performance varied in a more non-linear way with the amount of required depth movement. This was contrary to the performance that showed increased depth axis errors and decreased lateral axis errors proportional to required depth movement under the binocular condition.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Miall, R. C., Weir, D. J. & Stein, J. F. Manual tracking of visual targets by trained monkeys. *Behav. Brain Res.* **20**, 185–201 (1986).
2. van Roon, D., Caeyenberghs, K., Swinnen, S. P. & Smits-Engelsman, B. C. M. Development of feedforward control in a dynamic manual tracking task. *Child. Dev.* **79**, 852–865 (2008).
3. Ferguson, G. D., Duysens, J. & Smits-Engelsman, B. C. M. Children with developmental coordination disorder are deficient in a visuo-manual tracking task requiring predictive control. *Neuroscience* **286**, 13–26 (2015).
4. Bennett, S. J. & Barnes, G. R. Smooth ocular pursuit during the transient disappearance of an accelerating visual target: the role of reflexive and voluntary control. *Exp. Brain Res.* **175**, 1–10 (2006).
5. Miall, R. C. Task-dependent changes in visual feedback control: a frequency analysis of human manual tracking. *J. Mot. Behav.* **28**, 125–135 (1996).
6. Miall, R. C., Weir, D. J. & Stein, J. F. Intermittency in human manual tracking tasks. *J. Mot. Behav.* **25**, 53–63 (1993).
7. Pasalar, S., Roitman, A. V. & Ebner, T. J. Effects of speeds and force fields on submovements during circular manual tracking in humans. *Exp. Brain Res.* **163**, 214–225 (2005).
8. Roitman, A. V., Massaquoi, S. G., Takahashi, K. & Ebner, T. J. Kinematic analysis of manual tracking in monkeys: characterization of movement intermittencies during a circular tracking task. *J. Neurophysiol.* **91**, 901–911 (2004).
9. Viviani, P., Campadelli, P. & Mounoud, P. Visuo-manual pursuit tracking of human two-dimensional movements. *J. Exp. Psychol. Hum. Percept. Perform.* **13**, 62–78 (1987).
10. Engel, K. C. & Soechting, J. F. Manual tracking in two dimensions. *J. Neurophysiol.* **83**, 3483–3496 (2000).
11. Sugi, T., Nakamura, M., Ide, J. & Shibasaki, H. Modeling of motor control on manual tracking for developing a handmovement-compensation technique. *Artif. Life Robot.* **7**, 112–117 (2003).
12. Inoue, Y. & Sakaguchi, Y. Periodic change in phase relationship between target and hand motion during visuo-manual tracking task: behavioral evidence for intermittent control. *Hum. Mov. Sci.* **33**, 211–226 (2014).
13. Boyer, É. O., Bevilacqua, F., Susini, P. & Hanneton, S. Investigating three types of continuous auditory feedback in visuo-manual tracking. *Exp. Brain Res.* **235**, 691–701 (2017).
14. Mrotek, L. A., Gielen, C. C. A. M. & Flanders, M. Manual tracking in three dimensions. *Exp. Brain Res.* **171**, 99–115 (2006).
15. Parker, A. J. Binocular depth perception and the cerebral cortex. *Nat. Rev. Neurosci.* **8**, 379–391 (2007).
16. Skrandies, W. Assessment of depth perception using psychophysical thresholds and stereoscopically evoked brain activity. *Doc. Ophthalmol.* **119**, 209–216 (2009).
17. Qian, N. Binocular disparity and the perception of depth. *Neuron* **18**, 359–368 (1997).
18. Ding, J. & Levi, D. M. A unified model for binocular fusion and depth perception. *Vis. Res.* **180**, 11–36 (2021).
19. Melmoth, D. R. & Grant, S. Advantages of binocular vision for the control of reaching and grasping. *Exp. Brain Res.* **171**, 371–388 (2006).
20. Ban, H., Preston, T. J., Meeson, A. & Welchman, A. E. The integration of motion and disparity cues to depth in dorsal visual cortex. *Nat. Neurosci.* **15**, 636–643 (2012).
21. Nadler, J. W., Angelaki, D. E. & DeAngelis, G. C. A neural representation of depth from motion parallax in macaque visual cortex. *Nature* **452**, 642–645 (2008).
22. Preston, T. J., Kourtzi, Z. & Welchman, A. E. Adaptive estimation of three-dimensional structure in the human brain. *J. Neurosci.* **29**, 1688–1698 (2009).
23. Harris, J. M. & Wilcox, L. M. The role of monocularly visible regions in depth and surface perception. *Vision Res.* **49**, 2666–2685 (2009).
24. Ichikawa, M., Saida, S., Osa, A. & Munechika, K. Integration of binocular disparity and monocular cues at near threshold level. *Vis. Res.* **43**, 2439–2449 (2003).
25. Bachmann, D., Weichert, F. & Rinkenauer, G. Evaluation of the leap motion controller as a new contact-free pointing device. *Sensors* **15**, 214–233 (2014).
26. Ropelato, S., Menozzi, M., Michel, D. & Siegrist, M. Augmented reality microsurgery: a tool for training micromanipulations in ophthalmic surgery using augmented reality. *Simul. Healthc.* **15**, 122–127 (2020).
27. Liang, S.-F.M., Menozzi, M. & Huang, Y.-Y.R. A mechanism based on finger-sliding behavior for designing radial menus. *Int. J. Ind. Ergon.* **74**, 102869 (2019).
28. Choi, W., Lee, J., Yanagihara, N., Li, L. & Kim, J. Development of a quantitative evaluation system for visuo-motor control in three-dimensional virtual reality space. *Sci. Rep.* **8**, 13439 (2018).
29. Ikbal, M. S., Ramadoss, V. & Zoppi, M. Dynamic pose tracking performance evaluation of HTC vive virtual reality system. *IEEE Access* **9**, 3798–3815 (2021).
30. Servos, P., Goodale, M. A. & Jakobson, L. S. The role of binocular vision in prehension: a kinematic analysis. *Vision Res.* **32**, 1513–1521 (1992).
31. Servos, P. Distance estimation in the visual and visuomotor systems. *Exp. Brain Res.* **130**, 35–47 (2000).
32. Niechwiej-Szwedo, E., Cao, M. & Barnett-Cowan, M. Binocular viewing facilitates size constancy for grasping and manual estimation. *Vis. Basel* **6**, 2 (2022).
33. Coull, J., Weir, P. L., Tremblay, L., Weeks, D. J. & Elliott, D. Monocular and binocular vision in the control of goal-directed movement. *J. Mot. Behav.* **32**, 347–360 (2000).
34. Savelsbergh, G. J. & Whiting, H. T. The acquisition of catching under monocular and binocular conditions. *J. Mot. Behav.* **24**, 320–328 (1992).
35. Hayhoe, M., Gillam, B., Chajka, K. & Vecellio, E. The role of binocular vision in walking. *Vis. Neurosci.* **26**, 73–80 (2009).
36. Heinen, T. & Vinken, P. Monocular and binocular vision in the performance of a complex skill. *J. Sports Sci. Med.* **10**, 520–527 (2011).

37. Read, J. C. A., Begum, S. F., McDonald, A. & Trowbridge, J. The binocular advantage in visuomotor tasks involving tools. *Iperception* **4**, 101–110 (2013).
38. Maiello, G., Kwon, M. & Bex, P. J. Three-dimensional binocular eye-hand coordination in normal vision and with simulated visual impairment. *Exp. Brain Res.* **236**, 691–709 (2018).
39. Parker, P. R. L. *et al.* Distance estimation from monocular cues in an ethological visuomotor task. *eLife* **11**, 141 (2022).
40. Backus, B. T., Fleet, D. J., Parker, A. J. & Heeger, D. J. Human cortical activity correlates with stereoscopic depth perception. *J. Neurophysiol.* **86**, 2054–2068 (2001).
41. Smith, J. E. T. & Parker, A. J. Correlated structure of neuronal firing in macaque visual cortex limits information for binocular depth discrimination. *J. Neurophysiol.* **126**, 275–303 (2021).
42. Thomas, O. M., Cumming, B. G. & Parker, A. J. A specialization for relative disparity in V2. *Nat. Neurosci.* **5**, 472–478 (2002).
43. Alvarez, I., Hurley, S. A., Parker, A. J. & Bridge, H. Human primary visual cortex shows larger population receptive fields for binocular disparity-defined stimuli. *Brain Struct. Funct.* **226**, 2819–2838 (2021).
44. Bridge, H. Effects of cortical damage on binocular depth perception. *Philos. Trans. R. Soc. Lond. B Bio. Sci.* **371**, 1697 (2016).
45. Hadjidimitrakis, K., De Vitis, M., Ghodrati, M., Filippini, M. & Fattori, P. Anterior-posterior gradient in the integrated processing of forelimb movement direction and distance in macaque parietal cortex. *Cell Rep.* **41**, 111608 (2022).
46. Messier, J. & Kalaska, J. F. Covariation of primate dorsal premotor cell activity with direction and amplitude during a memorized-delay reaching task. *J. Neurophysiol.* **84**, 152–165 (2000).
47. Hadjidimitrakis, K., Ghodrati, M., Breveglieri, R., Rosa, M. G. P. & Fattori, P. Neural coding of action in three dimensions: task- and time-invariant reference frames for visuospatial and motor-related activity in parietal area V6A. *J. Comp. Neurol.* **528**, 3108–3122 (2020).
48. Churchland, M. M., Santhanam, G. & Shenoy, K. V. Preparatory activity in premotor and motor cortex reflects the speed of the upcoming reach. *J. Neurophysiol.* **96**, 3130–3146 (2006).
49. De Vitis, M. *et al.* The neglected medial part of macaque area PE: segregated processing of reach depth and direction. *Brain Struct. Funct.* **224**, 2537–2557 (2019).
50. Hocherman, S. & Levy, H. The role of feedback in manual tracking of visual targets. *Percept. Mot. Skills* **90**, 1235–1248 (2000).
51. Miall, R. C., Weir, D. J. & Stein, J. F. Planning of movement parameters in a visuo-motor tracking task. *Behav. Brain Res.* **27**, 1–8 (1988).
52. Choi, W., Li, L. & Lee, J. Characteristic of motor control in three-dimensional circular tracking movements during monocular vision. *Biomed Res. Int.* **2019**, 3867138 (2019).
53. Gibaldi, A. & Banks, M. S. Binocular eye movements are adapted to the natural environment. *J. Neurosci.* **39**, 2877–2888 (2019).
54. Parker, M. G., Weightman, A. P., Tyson, S. F., Abbott, B. & Mansell, W. Sensorimotor delays in tracking may be compensated by negative feedback control of motion-extrapolated position. *Exp. Brain Res.* **239**, 189–204 (2021).

Author contributions

Funding

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-48869-w.

Correspondence and requests for materials should be addressed to W.C. or J.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.