Intelligent Human Systems Integration (IHSI 2023), Vol. 69, 2023, 255–265
https://doi.org/10.54941/ahfe1002844

Multiphase Pointing Motion Model Based on Hand-Eye Bimodal Cooperative Behavior

Chenglong Zong, Xiaozhou Zhou, Jichen Han, and Haiyan Wang

School of Mechanical Engineering, Southeast University, Nanjing, China

ABSTRACT

Pointing, as the most common interaction behavior in 3D interactions, has become the basis and hotspot of natural human-computer interaction research. In this paper, hand and eye movement data of multiple participants in a typical pointing task were collected with a virtual reality experiment, and we further clarified the movements of the hand and eye in spatial and temporal properties and their cooperation during the whole task process. Our results showed that the movements of both the hand and eye in a pointing task can be divided into three stages according to their speed properties, namely, the preparation stage, ballistic stage and correction stage. Based on the verification of the phase division of hand and eye movements in the pointing task, we further clarified the phase division standards and the relationship between the duration of every pair of phases of hand and eye. Our research has great significance for further mining human natural pointing behavior and realizing more reliable and accurate human-computer interaction intention recognition.

Keywords: 3D user interface, Pointing, Human motor behavior, Hand-eye coordination

INTRODUCTION

Natural human-computer interaction in three-dimensional space is a hot research topic in the field of virtual reality (VR). Among them, pointing, as one of the most basic interaction tasks, has a critical influence on the VR user experience.

With deeper research, multi-phase movement of natural human behavior has been found in pointing (Argelaguet & Andujar, 2013). A widely accepted theory is the two-components theory proposed by Woodworth (1899), who argued that the motor process of completing pointing consists of two main phases: the ballistic stage of fast motion and the correction stage of slow motion.

This multi-phase feature reflects a neurological process: the human brain first makes overall planning of the motor behavior before executing the action, arranging large muscle groups with high strength to move successively with small muscle groups with high precision (König et al., 2009) to achieve speed-accuracy trade-off (van Galen et al., 1995). Thus, multistage features of pointing behavior are found in multiple modalities that depend on

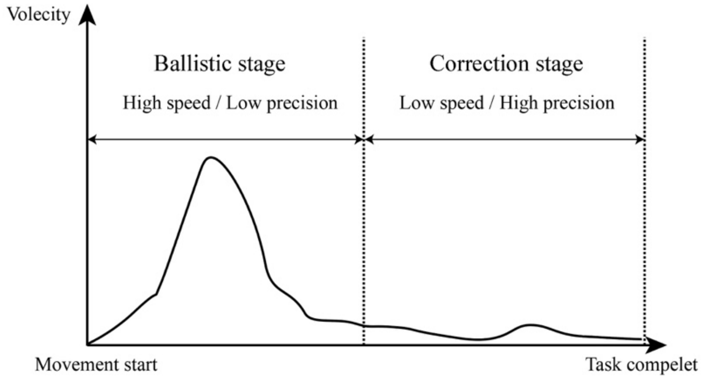

Figure 1: Velocity-task image in pointing proposed by the optimized pulse model.

muscle movements, such as the hand (Woodworth, 1899) and eye (Martin & Schovanec, 1997), which are most frequently used for pointing in VR.

The completion of pointing was not always the result of single-modal involvement, known as hand-eye coordination (Sailer et al., 1997). In pointing, the eye will perceive visual information, offer a target to hand, and timely adjust hand movement through feedback from visual information. The coordination effects of hand and eye during movement phases need to be further explored.

To better descript it, in this work, we build a multiphase behavioral model containing both hand and eye modalities based on collecting hand-eye bimodal data under a typical pointing task with VR. In this model, eye movement phases are divided into 3 phases through a multiphase division same with the hand rather than eye pattern to better emphasize the coordination between hand and eye movement.

The main contribution of our work is: 1. Verified the multiphase division methods can be used to divide eye movement phases in the pointing task. 2. The preparation stage was added to the common two-component model to temporally align the bimodal phases. 3. Based on the above, we built the bimodal multiphase movement model and confirmed it matches current research through a typical pointing experiment.

MULTIPHASE MOVEMENTS OF HAND AND EYE

The multiphase model typically explains multiple distinct behavioral phases that occur consecutively for a single modality, which is divided in the temporal dimension and usually manifests in differences in speed.

The two-component model of the hand has been extensively studied (Liu et al., 2009; Crossman & Goodeve, 1983), and the optimized impulse model (Meyer et al., 1988) proposed by Meyer et al. based on Woodworth (1899) is often used to explain the hand motion of a user performing a selection task. As shown in Figure 1, unlike Fitts' law, which directly orients the whole moving process, this model provides a more detailed description of the task process, which distinguishes the hand motion phase in the selection task into

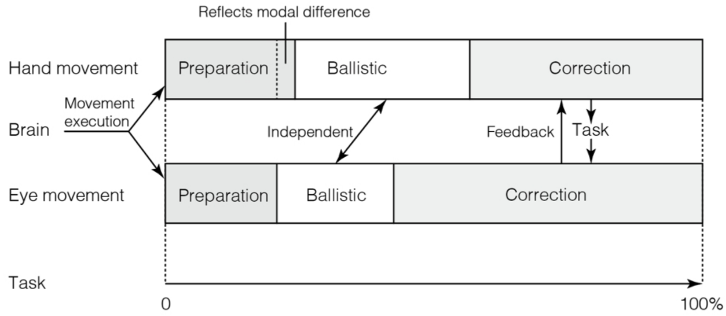

Figure 2: Movement phases and bimodal behavior in our proposed model.

a ballistic stage for fast movement with low precision and a slow correction stage with high precision.

The ballistic and correction stages differ in the mode of control of the motion. The ballistic stage is based on the execution of motion planning before the start of the motion, with less flexibility (Elliott et al., 2020), whereas the correction stage involves more feedback information.

Many studies have demonstrated that distance affects the duration of both the ballistic stage and the correction stage (Gan & Hoffmann, 1988; Hoffmann & Gan, 1988) and that greater distance extends the duration of both.

Eye both have multiple movement phases while perceiving visual information. Russo and Leclerc (1982) argued that in decision tasks, the motor phases of the eye can be divided into orientation, evaluation, and verification. While in a simple pointing task, since there is no complex decision-making process, it represents a cascade to perceive the target and a continuous fixation on the target to guide hand movement.

Some researchers have investigated the motor phases before the onset of movement in the same task for bimodality. Many studies (Biguer et al., 1982; Jeannerod, 1988) found that the hand started to move later than the eye started to move in the same task, with a time difference between 40-100ms. Thus, the time before the start of the hand-eye movement is always static in the same task. But other detailed phases haven't been researched along with its counterpart.

PROPOSED MODEL

We proposed a model that contains multiphase movements of the hand and eye. In our model, hand movement is divided into phases as in previous multiphase models. Besides, eye movements are also described with a multiphase method rather than divide into eye patterns such as cascade and fixation, to emphasize the potential relation between hand and eye movement phases.

We both divide the hand and eye motion process in pointing into three phases based on the speed characteristics, which are the preparation stage, ballistic stage, and correction stage. Figure 2 shows the division of our hand

and eye motion stages, and this division adds the preparation stage to the common two-component model.

The preparation stage describes the phase from the appearance of the task target to the start of the movement of the modality. As the hand and eye start moving at different times in the movement (Jeannerod, 1988), the preparation stage is critical to make the two aligned in the time sequence. As the human hand and eye receive planning simultaneously, the preparation stage duration of the eye should be smaller than the preparation stage duration of the hand, the difference is not affected by the distance and is fixed between 40-100ms, the same as the previous studies.

The ballistic stage represents the time from the start of acceleration to the maximum velocity of the modality and decelerates with anticipation after reaching the maximum velocity. The ballistic stage of the eye will be significantly faster than the ballistic stage of the hand due to the advantage of the eye in localization speed, and since there is no conscious involvement in these two phases, the two will execute relatively independent movements after receiving the motion signal, and the distance will have different effects on the two due to having different muscle movement characteristics.

In the correction stage, human achieves more precise but lower-speed movements. Note that although the hand and eye are relatively identical in terms of speed characteristics, where the hand is mainly responsible for outputting commands, while the eye continuously locks onto the target in this phase to perceive visual information. We named this stage of both hand and eye as the correction stage to facilitate the presentation and to express the close connection that both have in each stage.

EXPERIMENT

We designed a typical pointing experiment involving multiple participants. In the experiment, we recorded the participants' hand and eye movements, which were used to verify our methods. The data collected will be analyzed to evaluate the details of our modal.

A total of 15 participants participated in the experiment, gender-balanced (8 males, 7 females), aged from 22-27, all had a normal or corrected-to-normal vision at the time of participation, and all were right-handed. After completing the experiment, the participants were paid 50 RMB.

The experiment was conducted in a virtual environment with a 90 Hz video output frame rate and data acquisition frame rate (both hand and eye), and the virtual reality display used was the Varjo XR3, ensuring that the targets displayed are clear and reliable eye-tracking data. The controller we used was the HTC VIVE grip, which's positioning accuracy was within 1 mm. The experimental program was written in unity, and the experimental data acquisition was done in unity through C# scripts.

As the scene illustrated in Figure 3, participants used rays to select spherical objects with a diameter of 40 cm, at 25 m quickly with guaranteed accuracy. The generated location of the target stimuli is shown in Figure 3. All target stimuli will appear on a sphere with a radius of 25 m from the

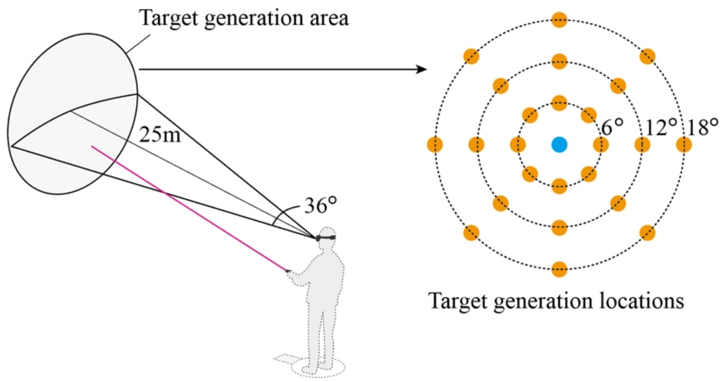

Figure 3: Experiment set.

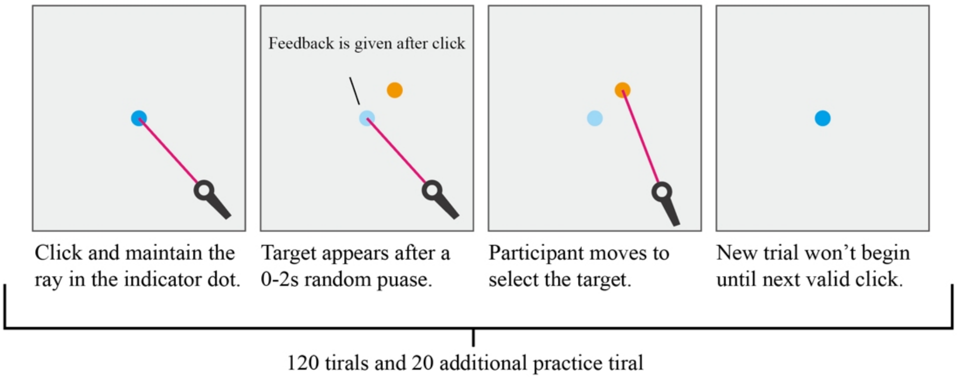

Figure 4: Procedure of a single trial.

localization origin, world-locked, and are offset in eight directions. The offset angles include three, 6, 12, and 18 degrees, which are angles that ensure that the participant does not experience spontaneous head movements. Directly 25 m away in front of the orientation origin is a 40 cm diameter indicator point, indicating that the participant needs to use this point as a starting point at each trial.

The participants were asked to wear the head-mounted display correctly, check the position and complete an eye-movement calibration using the built-in 9-point calibration mode.

In a single trial (see Figure 4), the participant was asked to click on the indicator dot and maintain the ray in the dot after the click to ensure that the starting point of the pointer was the same. Besides, this guides the participant to focus on the dot continuously, so that the eye would also look at the dot. Within a random time between 0–2 s after the click, a target stimulus will appear at one of the pre-set locations, and the participant needs to select the target stimulus as soon as possible within 3 s without failing. If it is not successfully selected at one time, the trial will be judged as unavailable to prevent the collected data from containing typical task processes.

The participants were required to complete 20 trials as practice. In the formal experiment, participants were required to complete 120 trials, i.e., each position was repeated five times, and participants could decide whether they needed to rest after each trial.

MOVEMENT PHASE DIVISION

To avoid high-frequency noise due to the natural jitter of the modalities, we used the method of Chen et al. (2015), where the horizontal and vertical directions of the hand and eye expressed in polar coordinates were passed through a 4th-order Butterworth low-pass filter at 10 Hz, respectively. The velocity and acceleration are then calculated from the processed information. Our movement phase division method is modified from Deng et al. (2019), which we judge as the dividing point of the movement phase when one of the following conditions is satisfied.

1) A smaller velocity (0.1 times the maximum velocity for the hand, 0.05 times the maximum velocity for the eye) and a negative acceleration. The different velocity set is based on different rotate speed of the bimodal.
2) The velocity is taken to a minimal point (acceleration past zero, and the acceleration changes from negative to positive), which means that the direction of muscle force changes.

We use the dividing point closest to the maximum velocity before the maximum velocity point as the dividing point between the preparation stage and the ballistic stage and the first dividing point after the maximum velocity point as the dividing point between the ballistic stage and the correction stage.

PHASE DIVISION RESULTS

A total of 15 participants participated in the experiment, and one participant was excluded due to a high error rate (45.2%). 1680 trials were obtained, and the average completion rate of the trials was 76.25%. Since the incomplete trials did not have a complete typical task flow, this part of the data was discarded. Among the remaining 1281 valid data, 1218 trials could complete the division completely, accounting for 95.08% of all valid data. Only the successfully divided data were further analyzed.

A typical trial in which we completed the division is shown in Figure 5 and presented as a velocity-task image. In this typical trial, both hand and eye data can complete a three-stage division.

The average velocity-task images of the hand and eye after superimposing all trial times at different distances are shown in Figure 6, and the hand images all have obvious three-stage characteristics, confirming the existence of hand motion behavior stages. The curve of eye is not smooth, but it also shows a concentrated area of velocity peaks, and the overall trend of the task process is rising and then falling.

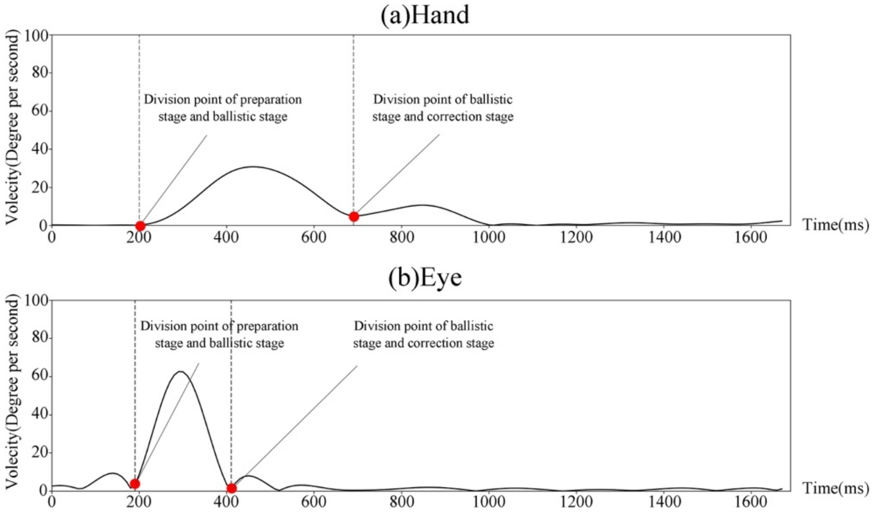

Figure 5: Velocity-time images of the hand and eye in a typical trial and the division of motion phases.

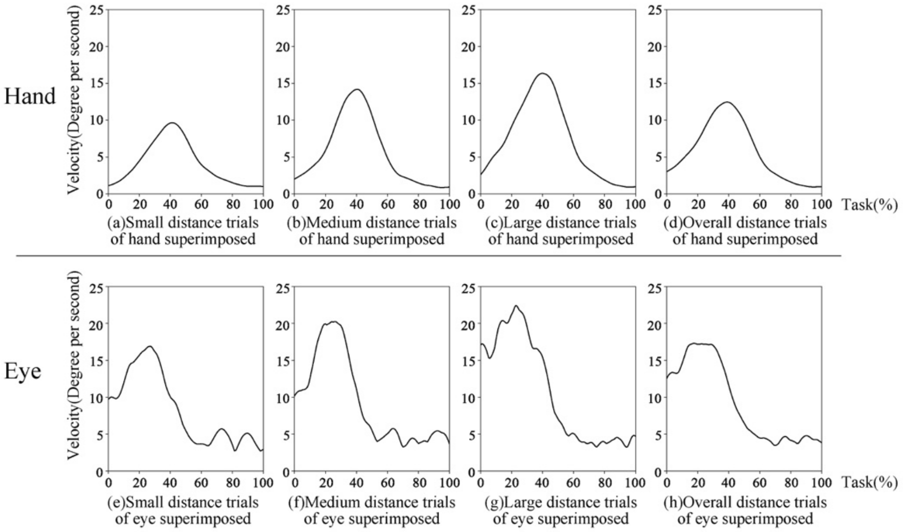

Figure 6: Hand and eye movement velocities for distances/overall trials superimposed with the task process.

COMPARISON OF THE PHASES

Based on the successfully divided data, we compared the quantitative relationship of the duration of the hand-eye preparation stage, to evaluate our model.

As seen in Figure 7(a), the average duration of the preparation stage of the hand 0.223s is longer than the average duration of the preparation stage of the eye 0.152s, and it behaves the same at different distances. ANOVA results showed that distance has a significant effect on the duration of both hands

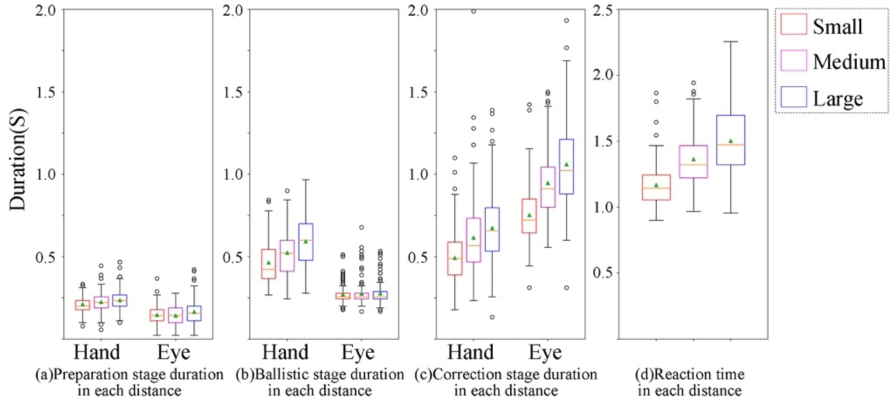

Figure 7: Duration of each phase of hand and eye as influenced by distance.

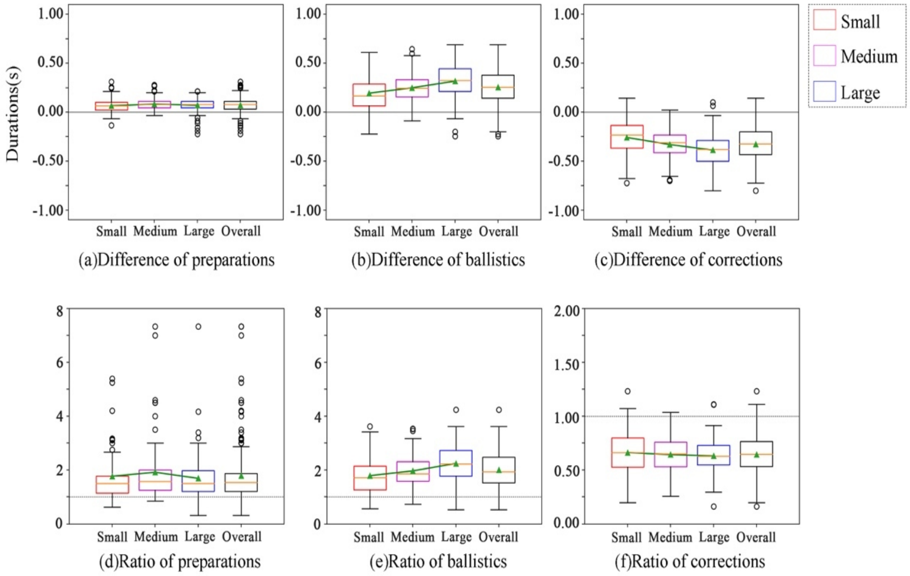

Figure 8: Comparison of hand and eye phase durations.

(F (2, 1215) = 8.64, p = 0.00) and eye (F (2, 1215) = 9.08, p = 0.00) in this phase, the preparation stage duration of hand increase with distance, but the eye's keep stead between small and medium distance (F (1, 790) = 0.71, p = 0.40), and significantly increased in large distance trials compared with the two others (F (1, 805) = 82.71, p = 0.00), (F (1, 835) = 10.33, p = 0.001). As shown in Figure 8, the average difference between the hand and eye preparation stages was 71.5Ms, and the distance did not have a significant effect on this value (F (2, 1215) = 2.42, p = 0.089). The mean ratio between hand and eye preparation stage durations was 1.797, and again, the distance did not have a significant effect on this value (F (2, 1215) = 0.90, p = 0.41).

The duration distribution of the hand and eye ballistic stage is shown in Figure 7(b), and the average duration of the ballistic stage of the hand 0.549s is greater than that of the eye 0.274s, which shows the same trend at different distances. As shown in Figure 8, the distance had different effects on the duration of the hand and eye in this phase, and the duration of the hand in this phase significantly increased with distance ($F(2, 1215) = 24.02$, $p = 0.00$), while the duration of the eye in this phase was not significantly affected by the distance ($F(2, 1215) = 0.18$, $p = 0.84$). Both the difference ($F(2, 1215) = 20.96$, $p = 0.00$) and the ratio ($F(2, 1215) = 17.07$, $p = 0.00$) of hand and eye durations in this phase increased with distance, the mean value of the difference was 0.255s, and the mean value of the ratio was 2.010.

The distribution of the duration of the correction stage of the hand and eye is shown in Figure 7(c), and the average duration of the correction stage of the hand 0.579s is smaller than the average duration of the correction stage of the eye 0.926s. The distance has a significant effect on the duration of both the correction stage of the hand ($F(2, 1215)= 28.61$, $p = 0.00$) and the eye ($F(2, 1215) = 68.55$, $p = 0.00$), and the duration of the correction stage of both the hand and the eye increases as the distance increases. As shown in Figure 8, the mean value of the difference between the correction stage of the hand and eye was −0.346s, and the absolute value of this difference was significantly affected by the distance ($F(2, 1215) = 22.69$, $p = 0.00$) and increased with distance. The mean value of the ratio between the correction stage of the hand and eye was 0.623, and distance did not significantly affect this value ($F(2, 1215) = 1.16$, $p = 0.31$).

The average maximum angular shift of the eye when the eye entered the correction stage until the trial was completed was 3.26 degrees across trials, meaning that the jitter of the eye's real-time gaze point after entering the correction stage was generally under 3.26 degrees.

DISCUSSION

The results show that we successfully delineated the hand and eye movement phases simultaneously during the behavior of the whole pointing task flow, and the data indicate that both modalities have multistage motion characteristics in this task. All three phases of both hand and eye modalities conformed to our expected velocity characteristics, showing small velocity in the preparation stage, peak velocity in the ballistic stage, and steady movement in the correction stage.

We independently analyzed the temporal relationships in the three phases of hand and eye and analyzed the effect of distance on the duration of each phase to infer the trend of synergistic changes in the phases of the hand and eye.

We observed that the duration of the preparation stage was generally longer in the hand than in the eye, with a mean value of approximately 70ms for this difference, and that this difference was not affected by distance, suggesting that the difference in duration between hand and eye in this phase is determined by fixed physiological factors and that the hand has a longer response and motor latency than the eye before receiving a movement execution. Same with previous studies. We also found that the duration of the hand and eye preparation stage simultaneously increased with the distance requirement in the pointing task reflects the distance was influenced by some hand-eye shared cognitive link within that phase.

After the eye and hand entered the ballistic stage one after another, the eye velocity was significantly better than the hand velocity, and the duration of the eye in this phase was not affected by the distance, while the duration of the ballistic stage of the hand increased with increasing distance. Furthermore, in the comparison of the quantitative relationship between hand and eye ballistic stage durations, it was found that the difference and ratio of both increased with time. Since changing the task requirements did not have similar effects on both, we consider the hand and eye movements in this phase to be independent, and consistent with our description.

The eye always entered the correction stage earlier than the hand entered the correction stage, and because both terminated at task completion, the eye's correction stage duration was always longer than the hand's correction stage duration. We verified that the eye enters the correction stage with less movement and maintains a fixation. The maximum offset of 3.26 degrees on average allows the target to be maintained in the human clear field of view, which is consistent with our definition of the characteristics of this phase. We found that the absolute value of the difference between the two increased with distance and unexpectedly found that distance did not have a significant effect on the ratio of the two, which had a mean value of approximately 0.6. We speculate that the ratio reflects the corrective effect in which the eye affects the hand with visual feedback in this certain task.

CONCLUSION

In this paper, human behavioral data were collected through pointing experiments in virtual reality to explore the movement phases of hand-eye bimodality in the pointing task. We verified eye movement in pointing can be divided with multiphase methods. The research results proved our model through validate the phase division of hand-eye behavior in pointing task and further clarify the phase division basis of bimodality and the relationship between the duration of each phase of bimodality. This research has provided a new framework for further deepening natural human pointing behavior and we will focus on improve current interaction technologies with its guide.

ACKNOWLEDGMENT

This work is supported in part by the National Natural Science Foundation of China (No. 71901061 and 52275238).

REFERENCES

Argelaguet, F., & Andujar, C. (2013). A survey of 3D object selection techniques for virtual environments. *Computers & Graphics*, *37*(3), 121–136.

Biguer, B., Jeannerod, M., & Prablanc, C. (1982). The coordination of eye, head, and arm movements during reaching at a single visual target. *Experimental brain research*, *46*(2), 301–304.

Chen, Y., Hoffmann, E. R., & Goonetilleke, R. S. (2015). Structure of hand/mouse movements. *IEEE Transactions on Human-Machine Systems*, *45*(6), 790–798.

Crossman, E. R. F., & Goodeve, P. J. (1983). Feedback control of hand-movement and Fitts' law. *The Quarterly Journal of Experimental Psychology Section A*, *35*(2), 251–278.

Deng, C. L., Geng, P., Hu, Y. F., & Kuai, S. G. (2019). Beyond Fitts's law: a three-phase model predicts movement time to position an object in an immersive 3D virtual environment. *Human Factors*, *61*(6), 879–894.

Elliott, D., Lyons, J., Hayes, S. J., Burkitt, J. J., Hansen, S., Grierson, L. E., ... & Bennett, S. J. (2020). The multiple process model of goal-directed aiming/reaching: insights on limb control from various special populations. *Experimental Brain Research*, *238*(12), 2685–2699.

Gan, K. C., & Hoffmann, E. R. (1988). Geometrical conditions for ballistic and visually controlled movements. *Ergonomics*, *31*(5), 829–839.

Hoffmann, E. R., & Gan, K. C. (1988). Directional ballistic movement with transported mass. *Ergonomics*, *31*(5), 841–856.

Jeannerod, M. (1988). The neural and behavioural organization of goal-directed movements. Clarendon Press/Oxford University Press.

König, W. A., Gerken, J., Dierdorf, S., & Reiterer, H. (2009, August). Adaptive pointing–design and evaluation of a precision enhancing technique for absolute pointing devices. In *IFIP Conference on Human-Computer Interaction* (pp. 658–671). Springer, Berlin, Heidelberg.

Liu, L., van Liere, R., Nieuwenhuizen, C., & Martens, J. B. (2009, March). Comparing aimed movements in the real world and in virtual reality. In *2009 IEEE Virtual Reality Conference* (pp. 219–222). IEEE.

Martin, C. F., & Schovanec, L. (1997, December). A control model of eye movement. In *Proceedings of the 36th IEEE Conference on Decision and Control* (Vol. 2, pp. 1135–1139). IEEE.

Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., & Keith Smith, J. E. (1988). Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological review*, *95*(3), 340.

Russo, J. E., & Leclerc, F. (1994). An eye-fixation analysis of choice processes for consumer nondurables. *Journal of consumer research*, *21*(2), 274–290.

Sailer, U., Eggert, T., Ditterich, J., & Straube, A. (2000). Spatial and temporal aspects of eye-hand coordination across different tasks. *Experimental Brain Research*, *134*(2), 163–173.

van Galen, G. P., & de Jong, W. P. (1995). Fitts' law as the outcome of a dynamic noise filtering model of motor control. *Human movement science*, *14*(4-5), 539–571.

Woodworth, R. S. (1899). Accuracy of voluntary movement. *The Psychological Review: Monograph Supplements*, *3*(3), i.