

• Article •

# Temporal continuity of visual attention for future gaze prediction in immersive virtual reality

Zhiming HU<sup>1</sup>, Sheng LI<sup>1,2\*</sup>, Meng GAI<sup>1,2</sup>

1. Department of Computer Science and Technology, School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, China

2. Beijing Engineering Technology Research Center of Virtual Simulation and Visualization, Peking University, Beijing 100871, China

\* Corresponding author, lisheng@pku.edu.cn

Received: 30 October 2019 Accepted: 9 January 2020

Supported by the National Key R&D Program of China (2017YFB0203000) and National Natural Science Foundation of China (61632003; 61661146002; 61631001).

**Citation:** Zhiming HU, Sheng LI, Meng GAI. Temporal continuity of visual attention for future gaze prediction in immersive virtual reality. *Virtual Reality & Intelligent Hardware*, XXXX, XX(XX): 1–11  
DOI: 10.1016/j.vrih.XXXX

**Abstract Background** Eye tracking technology is receiving increased attention in the field of virtual reality. Specifically, future gaze prediction is crucial in pre-computation for many applications such as gaze-contingent rendering, advertisement placement, and content-based design. To explore future gaze prediction, it is necessary to analyze the temporal continuity of visual attention in immersive virtual reality. **Methods** In this paper, the concept of temporal continuity of visual attention is presented. Subsequently, an autocorrelation function method is proposed to evaluate the temporal continuity. Thereafter, the temporal continuity is analyzed in both free-viewing and task-oriented conditions. **Results** Specifically, in free-viewing conditions, the analysis of a free-viewing gaze dataset indicates that the temporal continuity performs well only within a short time interval. A task-oriented game scene condition was created and conducted to collect users' gaze data. An analysis of the collected gaze data finds the temporal continuity has a similar performance with that of the free-viewing conditions. Temporal continuity can be applied to future gaze prediction and if it is good, users' current gaze positions can be directly utilized to predict their gaze positions in the future. **Conclusions** The current gaze's future prediction performances are further evaluated in both free-viewing and task-oriented conditions and discover that the current gaze can be efficiently applied to the task of short-term future gaze prediction. The task of long-term gaze prediction still remains to be explored.

**Keywords** Temporal continuity; Visual attention; Autocorrelation analysis; Gaze prediction; Virtual reality

## 1 Introduction

Eye tracking technology aims at tracking users' gaze positions and has many important applications in the area of virtual reality (VR) including eye movement-based interaction<sup>[1,2]</sup>, gaze-contingent rendering<sup>[3,4]</sup>, gaze behavior analysis<sup>[5-7]</sup>, and foveated imaging<sup>[8]</sup>. Eye tracking methods can be classified into real-time and future gaze prediction methods. Currently, the most common solution for real-time gaze prediction is based on eye trackers. An eye tracker is hardware integrated into head mounted devices<sup>[9]</sup> (HMDs). In addition to eye trackers, a software-based solution is also proposed for real-time gaze prediction in virtual

reality<sup>[6]</sup>. However, compared with real-time gaze prediction, there is limited work on future gaze prediction. Future gaze prediction is crucial in pre-computation for many applications such as gaze-contingent rendering, advertisement placement, and content-based recommendation. To explore the topic of future gaze prediction, there is a need to analyze the temporal continuity of visual attention.

In this paper, the concept of temporal continuity of visual attention in immersive virtual reality is presented. Temporal continuity refers to the continuity and consistency of users' on-screen gaze position sequences. An autocorrelation function (ACF) was utilized to evaluate the temporal continuity. As revealed in prior works<sup>[10,11]</sup>, there exist two mechanisms of visual attention: a top-down mechanism and a bottom-up mechanism. The temporal continuity of visual attention under the top-down mechanism may perform differently from the bottom-up mechanism. Therefore, temporal continuity under both mechanisms is analyzed independently. Specifically, free-viewing conditions (bottom-up mechanism) and task-oriented conditions (top-down mechanism) are explored.

**Free-viewing conditions:** In free-viewing conditions, an analysis of a free-viewing gaze dataset is performed<sup>[6]</sup> (Section 4) by calculating the ACF of users' gaze position sequences to evaluate the temporal continuity. It was found that the ACF only performs well within 100ms. The ACF deteriorates significantly with the increase of the time interval and becomes very small when the time interval is larger than 700ms.

**Task-oriented conditions:** To analyze the temporal continuity of visual attention in task-oriented conditions, a task-oriented game scene was created and a user study was conducted to collect 19 players' gaze data (Section 5). Temporal continuity was analyzed using the collected gaze data. The ACF in task-oriented conditions has similar characteristics with that of the free-viewing conditions.

**Future gaze prediction:** The temporal continuity of visual attention for the task of future gaze prediction was further applied. If the temporal continuity is good, users' current gaze positions can be directly employed to predict their gaze positions in the future. It was found that, in both free-viewing (Section 4.3) and task-oriented conditions (Section 5.3), current gaze only performs well for short-term gaze prediction and cannot efficiently handle long-term gaze prediction.

Overall, the contributions of this study include

- (1) The concept of temporal continuity of visual attention in immersive virtual reality with a method to evaluate it.
- (2) The temporal continuity of visual attention in both free-viewing conditions and task-oriented conditions.
- (3) Application of temporal continuity to future gaze prediction and evaluation of its performance.

## 2 Related work

This section provides a brief overview of prior works on visual attention, the temporal characteristics of visual attention, and gaze prediction.

### 2.1 Visual attention

Analyzing human visual attention is an active area of vision research. Many prior works revealed that human visual attention is controlled by two mechanisms, bottom-up and top-down<sup>[10,11]</sup>. The bottom-up mechanism is fast and biases the attention towards the salient regions of the content, while the top-down mechanism is slow and it directs human visual attention to task-related objects. The two mechanisms are found to be independent<sup>[12]</sup>. In addition, the horizontal and vertical eye movements are found to behave differently<sup>[13]</sup>.

Human visual attention has also been studied in the field of virtual reality and has many applications. Sitzmann et al. found that there exists an equator bias when users are watching 360° images and they utilized this bias to adapt existing saliency predictors<sup>[14]</sup>. Hu et al. revealed a linear correlation between

users' gaze positions and their head rotation velocities, and they further employed users' head movements to predict their real-time gaze positions<sup>[6]</sup>. In this paper, the focus is on the temporal characteristics of visual attention and the application of future gaze prediction.

## 2.2 Temporal characteristics of visual attention

The temporal characteristics of visual attention have been studied by many researchers. Henderson focused on the temporal characteristics of visual attention during real-world scene perception<sup>[15]</sup>. He revealed that the average fixation duration during real-world scene viewing is around 330 ms, although there exists a large variability in this approximation. The length of fixation durations is found to be influenced by both the low-level features of the scene, such as luminance<sup>[16]</sup>, which impact bottom-up processing, and high-level features<sup>[17]</sup>, which influence top-down processing.

In the field of virtual reality, the temporal characteristics of visual attention have also been studied. Sitzmann et al. revealed that observers in virtual reality behave in two different modes: "attention" and "re-orientation"<sup>[14]</sup>. Attention mode refers to the condition when observers focus their attention on some regions, while re-orientation mode is the status when observers shift their attention. Hu et al. focused on the temporal characteristics of visual attention in free-viewing conditions<sup>[6]</sup>. They reported that saccades, which refer to fast eye movements, seldom occur in free-viewing conditions. In this paper, the focus is on the temporal continuity of visual attention in immersive virtual reality.

## 2.3 Gaze prediction

Gaze prediction, or visual saliency prediction, is a hot issue in the area of vision research and many gaze prediction methods have been proposed. Generally, most of the existing gaze prediction methods are based on bottom-up models<sup>[18,19]</sup>, which focus on low-level image features like intensity, color, and orientation, or top-down models<sup>[20,21]</sup>, which take high-level features such as specific tasks and context into consideration. In addition, with recent advances in deep learning, many deep learning-based gaze prediction methods have also been proposed<sup>[22]</sup>.

In the area of virtual reality, however, there is little work on gaze prediction. Sitzmann et al. focused on the saliency in 360° static images<sup>[14]</sup>. They conducted a study to collect users' eye tracking data in 360° images and proposed a method to predict saliency maps in virtual reality. Koulieris et al. focused on gaze prediction in a task-oriented video game<sup>[23]</sup>. They proposed a machine learning-based method to predict the object categories that users gazed at during game play. Hu et al. concentrated on real-time gaze prediction in virtual reality under free-viewing conditions<sup>[6]</sup>. They proposed an eye-head coordination model for real-time gaze prediction. In this paper, the feasibility of future gaze prediction in immersive virtual reality is explored.

# 3 Temporal continuity of visual attention

In this section, the concept of temporal continuity is clarified, including its definition, importance, and application. Subsequently, a method to evaluate the temporal continuity is presented.

## 3.1 The concept of temporal continuity

In this research, the temporal continuity of visual attention is defined in immersive virtual reality as the continuity and consistency of users' on-screen gaze position sequences. A Cartesian coordinate is used to describe users' on-screen gaze data by setting the origin to the center of the HMD screen, orienting the X-axis from left to right, and the Y-axis from bottom to top. As utilized in prior works<sup>[6,24]</sup>, users' horizontal

and vertical on-screen gaze positions using visual angle are measured, i.e., the angle between a user's line of sight and the normal direction of the HMD screen plane. For example, if a user fixates on the HMD screen center, his on-screen gaze position will be  $(0^\circ, 0^\circ)$ .

The temporal continuity of visual attention is very important for future gaze prediction. Currently, eye trackers are mainly designed to measure users' current gaze positions and cannot predict users' gaze positions in the future. For short-term future gaze prediction, there may be no need to develop new eye tracking technology. However, for long-term future gaze prediction, there is a necessity to propose accurate gaze prediction methods. Analyzing the temporal continuity of visual attention can help determine the time interval at which a gaze prediction method is needed.

If users' on-screen gaze positions have good temporal continuity, users' current gaze positions can be directly employed to predict their gaze positions in the future:

$$\begin{aligned}x_g(t_0 + \Delta t) &= x_g(t_0), \\y_g(t_0 + \Delta t) &= y_g(t_0),\end{aligned}\tag{1}$$

where  $x_g(t_0)$  and  $y_g(t_0)$  are the current horizontal and vertical gaze positions, respectively;  $t_0$  is the current time;  $x_g(t_0 + \Delta t)$  and  $y_g(t_0 + \Delta t)$  are the horizontal and vertical gaze positions in the future, respectively; and  $\Delta t$  is the time interval. One of the goals is to determine the value of  $\Delta t$ , for which the prediction performance of Equation 1 is considerable.

### 3.2 The evaluation of temporal continuity

If users' gaze positions have good temporal continuity, their current gaze positions will be highly correlated with their gaze positions in the near future. In other words, users' gaze position sequences have autocorrelation. Therefore, to evaluate the temporal continuity, the ACF of users' gaze position sequences is calculated by estimating the correlation between the gaze position sequence and a delayed copy of the sequence. Specifically, the ACF of users' horizontal and vertical gaze position sequences are calculated using the formula proposed in the work done by Box et al.<sup>[25]</sup>:

$$r_k = \frac{1}{c_0 T} \sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y}),\tag{2}$$

where  $r_k$  is the autocorrelation function of  $y_t$  and it lies in the range  $[-1, 1]$ , with  $-1$  indicating perfect anti-correlation and  $1$  indicating perfect correlation. Generally, an absolute value for  $r_k$  of  $0.1$  is classified as small,  $0.3$  is medium,  $0.5$  is strong, and  $0.7$  is high<sup>[26,27]</sup>.  $y_t$  is the horizontal or vertical gaze position sequence whose autocorrelation is analyzed;  $c_0$  is the variance of  $y_t$ ;  $T$  is the number of gaze data in  $y_t$ , i.e., the sequence length; and  $\bar{y}$  is the mean of  $y_t$ .  $y_{t+k}$  is a delayed copy of  $y_t$  and  $k$  is the lag between sequence  $y_t$  and sequence  $y_{t+k}$ . For example, if  $k = 10$ , sequence  $y_t$  is  $(y_1, y_2, y_3, \dots, y_T)$  and sequence  $y_{t+k}$  is  $(y_{11}, y_{12}, y_{13}, \dots, y_T)$ . In this calculation,  $k = 0, 1, 2, \dots, 100$  to calculate the autocorrelation. As the gaze position sequence analyzed in this research is sampled at every 10ms, the time intervals between  $y_t$  and  $y_{t+k}$  for  $k = 0, 1, 2, \dots, 100$  are  $0, 10, 20, \dots, 1000$ ms.

## 4 Free-viewing conditions

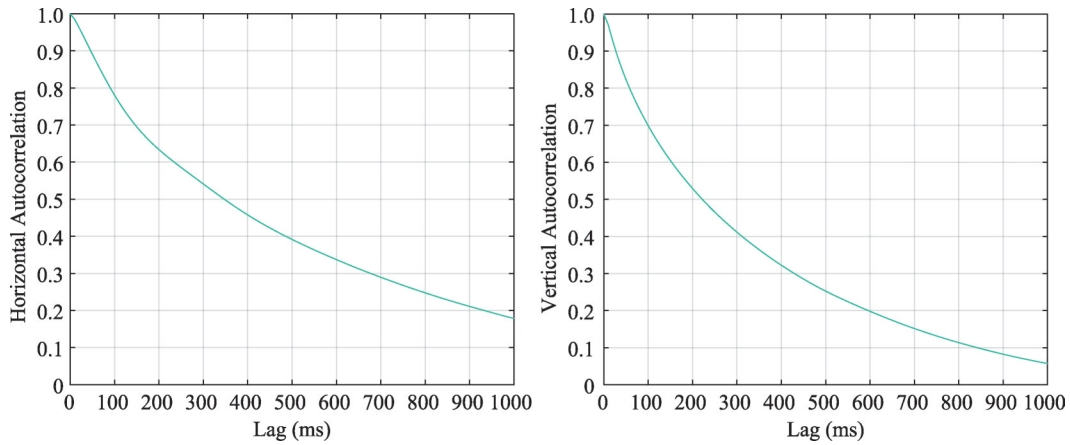
In this section, the temporal continuity of visual attention in free-viewing conditions is analyzed. Specifically, the analysis of a free-viewing gaze dataset<sup>[6]</sup> is performed. The autocorrelation function of the gaze position sequence is first calculated to assess the temporal continuity. Then the temporal continuity is applied to future gaze prediction to evaluate its performance.

## 4.1 Gaze data

Hu et al. recently studied human gaze behaviors under free-viewing conditions in immersive virtual reality and built a large eye tracking dataset<sup>[6]</sup>, which contained 60 participants' free-viewing gaze data in 7 static virtual scenes. During their data collection process, each participant was asked to explore 2 scenes in 2 lighting conditions and thus the dataset contains 240 pieces of data in total. Each piece of data contains a participant's continuous exploration data in a scene and it can be utilized to analyze the temporal continuity of visual attention. This dataset contains over 4000000 gaze positions and therefore is large enough to be employed for gaze behavior analysis. Therefore, for simplicity, temporal continuity is directly analyzed based on this dataset.

## 4.2 Temporal continuity evaluation

To evaluate temporal continuity, Equation 2 is employed to calculate the autocorrelation function of users' gaze position sequences. Since users' horizontal and vertical gaze behaviors are different<sup>[13]</sup>, the ACFs of the horizontal and vertical gaze position sequences are estimated. Since there are 240 pieces of data in total, the ACF of each piece of data is calculated first. Then the mean of the 240 ACFs is determined and the mean ACF is utilized as the ACF of users' gaze position sequences in free-viewing conditions. Figure 1 illustrates the horizontal and vertical autocorrelation functions, which show that both the horizontal and vertical ACFs decrease with the increase of lag time. Within the range of 100ms, the horizontal ACF is larger than 0.75 and the vertical ACF is larger than 0.7. When the lag increases to 400ms, the values of both the horizontal and vertical ACFs decrease significantly. The horizontal ACF decreases to around 0.45 and the vertical ACF decreases to around 0.3. When the lag is larger than 700ms, the values of the ACFs become very small. The horizontal ACF decreases to less than 0.3 and the vertical ACF decreases to less than 0.15.



**Figure 1** The horizontal (left) and vertical (right) autocorrelation functions of users' gaze position sequences in free-viewing conditions.

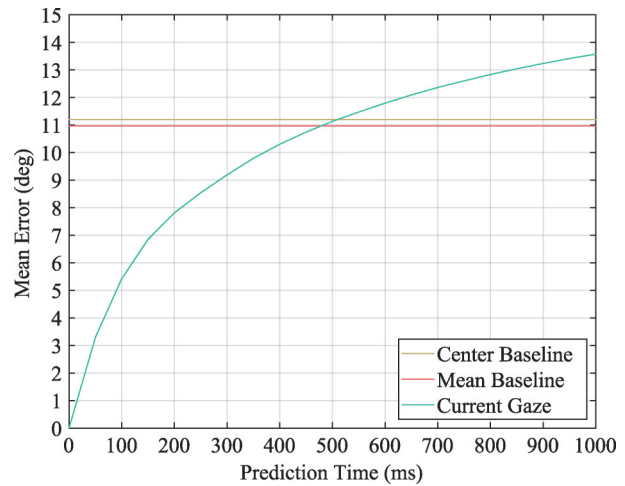
The above analysis reveals the characteristics of the temporal continuity of visual attention in free-viewing conditions. It can be concluded that the temporal continuity performs well within a short time (100ms or less), decreases significantly when the time interval increases, and becomes very weak after a long time (700ms or more).

## 4.3 Future gaze prediction

An important application of the temporal continuity of visual attention is future gaze prediction. If the

temporal continuity is good, users' current gaze positions can be directly utilized to predict their gaze positions in the future (Equation 1). To better evaluate the performance of gaze position prediction, an evaluation metric was set. Specifically, the **angular distance** was utilized as the evaluation metric between the ground truth and the predicted gaze position, i.e., the angle between the user's ground truth line of sight and the predicted line of sight. The smaller the angular distance, the smaller the prediction error and the better the performance. In addition, two baselines that were proposed in Hu et al.'s work<sup>[6]</sup>, the screen center (**Center Baseline**), which is  $(0^\circ, 0^\circ)$ , and the mean of all the gaze positions (**Mean Baseline**), which is  $(0.13^\circ, -2.32^\circ)$ , were used as this study's baselines.

To evaluate temporal continuity's performance on future gaze prediction the study employed current gaze positions, Center baseline, and Mean baseline to predict gaze positions in the future at 50, 100, 150, ..., 1000ms, and calculate their mean prediction errors (mean angular distances). Figure 2 illustrates the prediction results. The Center and Mean baselines retain the same performance at different prediction times because they are constant. Current gaze indicates good performance within 100ms. Its prediction performance deteriorates significantly with the increase of prediction time. At the prediction time of 600ms, it performs even worse than the baselines. The above results indicate that, in free-viewing conditions, temporal continuity can significantly improve the performance of short-term gaze prediction but it cannot efficiently handle long-term gaze prediction. The left of Figure 3 illustrates a user's gaze trajectory in free-viewing conditions.



**Figure 2** Current gaze and the baselines' future prediction performances in free-viewing conditions.



**Figure 3** Users' gaze trajectories in free-viewing conditions (left) and task-oriented conditions (right). The green cross refers to users' current gaze positions; blue indicates users' gaze positions in future 100ms; yellow denotes gaze positions in future 400ms; red represents gaze positions in future 700ms.

## 5 Task-oriented conditions

In this section, the temporal continuity of visual attention in task-oriented conditions is analyzed. Human

visual attention in task-oriented conditions is different from that in free-viewing conditions in that user's visual attention is influenced by the specific tasks assigned to them. In virtual reality, games are very common task-oriented applications. Therefore, to analyze the temporal continuity, a task-oriented game scene was created to collect users' gaze data. Autocorrelation analysis of the data was performed to evaluate the temporal continuity and also to measure temporal continuity's future gaze prediction performance.

## 5.1 Gaze data

To explore task-oriented conditions, a task-oriented game was created and a user study was conducted to collect users' gaze data.

**Stimuli:** A game scene using the Unity game engine with randomly placed animals such as ibexes and deer. The animals are dynamic and their movements are controlled by their own animations. The animals' paths are controlled using a Unity script developed for this study, which allows the animals to wander in the scene in a random manner. The animals are utilized as the targets in the game. The snapshot of this game is demonstrated in the left of Figure 4.



Figure 4 Left: The task-oriented game scene; Right: The experimental setup.

**Participants:** In total, 19 players (13 males, 6 females, ages 18–28) participated in the user study. Each participant reported normal or corrected-to-normal vision. The eye tracker was calibrated for each player before he/she started the game.

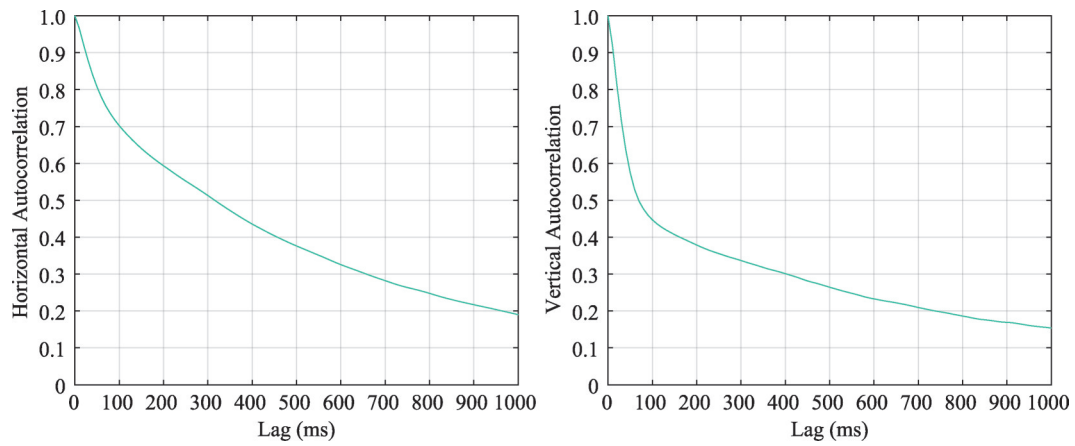
**System details:** In this user study, an HTC Vive was used as the HMD to display the game and a Vive controller was utilized for user interaction. A 7 invensun VR eye tracker, with a sampling frequency of 100Hz and accuracy of  $0.5^\circ$ , was used to collect users' gaze data. The game scene is created using the Unity game engine. The CPU and GPU of the platform are an Intel(R) Core (TM) i7-8700 @ 3.20GHz and an NVIDIA GeForce RTX 2080 Ti, respectively. The snapshot of the experimental setup is demonstrated in the right of Figure 4.

**Procedure:** The players were given a Vive controller to help teleport themselves into the scene. They were given a wand, which was controlled by the Vive controller, to hit the targets in the game, i.e., the animals. The more targets they hit, the higher their scores. A target will disappear if it is hit. Before starting the game, each player was given at least 3 minutes to become familiar with the experimental system. The players were asked to engage in the game for at least 2 minutes, and during the game, their gaze data was collected for later analysis.

**Gaze data:** 19 players participated in the game and thus there are 19 pieces of data in total. Each piece of data contains at least 12000 gaze positions, for a total of approximately 300000 gaze positions.

## 5.2 Temporal continuity evaluation

To evaluate temporal continuity, an autocorrelation analysis of the gaze data collected in the user study was performed by utilizing Equation 2. The ACFs of the 19 pieces of data and then the mean ACF were calculated. The mean ACF was utilized as the ACF of users' gaze position sequences in task-oriented conditions. As illustrated in Figure 5, similar to the ACFs in free-viewing conditions, the ACFs in task-oriented conditions decrease with the increase of lag time. In both the horizontal and vertical directions, the values of the ACFs are relatively high within 100ms; the ACFs deteriorate significantly with the increase of lag; and when the lag is larger than 700ms, the values of the ACFs become very small (horizontal ACF  $<0.3$ , vertical ACF  $<0.25$ ).

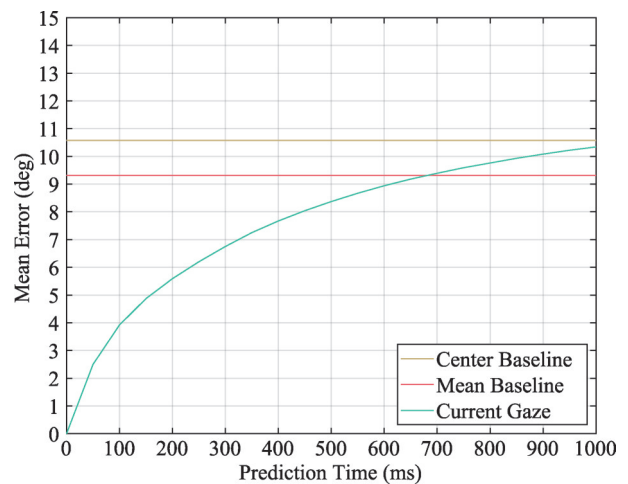


**Figure 5** The horizontal (left) and vertical (right) autocorrelation functions of users' gaze position sequences in task-oriented conditions.

This analysis reveals the characteristics of the temporal continuity of visual attention in task-oriented conditions. The temporal continuity decreases with the increase of the time interval. It performs well within a short time interval (100ms or less) and seriously deteriorates if the time interval is very large (700ms or more).

## 5.3 Future gaze prediction

The future gaze prediction performance of current gaze positions was also evaluated, based on the gaze data collected in the game. The angular distance is utilized as the evaluation metric and the Center and Mean baselines were employed as the study's baselines. In this case, Mean baseline refers to the mean of all the gaze data, which is  $(4.88^\circ, 0.97^\circ)$ . Current gaze, Center baseline, and Mean baseline's gaze prediction performances were calculated for 50, 100, 150, ..., 1000ms in the future. Figure 6 illustrates the mean prediction errors of current gaze and the baselines, and illustrates that the performances of the Center and Mean baselines are constant. Within 100ms, current gaze retains high accuracy. However, with the increase of prediction time, the accuracy of current gaze deteriorates



**Figure 6** Current gaze and the baselines' future prediction performances in task-oriented conditions.



significantly. Current gaze performs worse than Mean baseline when the prediction time is larger than 700ms. These results indicate that the temporal continuity is only effective for short-term gaze prediction (100ms or less). A user's gaze trajectory in task-oriented conditions is illustrated in the right of Figure 3.

## 6 Conclusions, limitations, and future work

In this paper, the concept of temporal continuity of visual attention in immersive virtual reality is presented and a novel analysis of the temporal continuity is discussed. The temporal continuity in both free-viewing and task-oriented conditions is evaluated by calculating autocorrelation functions of users' gaze position sequences. In free-viewing conditions, autocorrelation analysis of a free-viewing gaze dataset was performed and it was discovered that the autocorrelation performs well only when the lag is small. In task-oriented conditions, a game scene was created and a user study conducted to collect users' gaze data. The autocorrelation functions were calculated based on the collected gaze data, showing that the temporal continuity for task-oriented conditions is similar to that under free-viewing conditions. The temporal continuity was further applied to the task of future gaze prediction, i.e., utilizing current gaze positions to predict gaze positions in the future. Future gaze prediction is vital in pre-computation for many applications such as gaze-contingent rendering, advertisement placement, and content-based recommendation. The gaze prediction performances reveal that, in both free-viewing and task-oriented conditions, temporal continuity can only efficiently facilitate short-term gaze prediction. With the increase of prediction time, the efficiency of temporal continuity deteriorates significantly. The task of long-term gaze prediction remains to be explored.

There are some limitations in this work. First, the mechanism of the temporal continuity of visual attention has not been explored thoroughly in this analysis. The mechanism of temporal continuity is intricate and the focus of this paper was only on the characteristics of temporal continuity. Exploring the mechanism of temporal continuity is an interesting avenue for future work. Second, when analyzing temporal continuity in task-oriented conditions, only a VR game scene was taken into consideration so the results might have a bias to the recorded data. The temporal continuity of visual attention in other VR games, such as multi-party games and other VR applications such as VR shopping, VR training, and VR education remain to be explored. Third, the influence of sound on the temporal continuity of visual attention is not considered in this work. In this analysis, both the free-viewing and task-oriented gaze data are collected from silent scenes. However, temporal continuity may be influenced if sound exists in the scenes. Therefore, considering the influence of sound on temporal continuity may further improve this work.

## References

- 1 Duchowski A T. Gaze-based interaction: a 30 year retrospective. *Computers & Graphics*, 2018, 73, 59–69  
DOI:10.1016/j.cag.2018.04.002
- 2 Mardanbegi D, Mayer B, Pfeuffer K, Jalaliniya S, Gellersen H, Perzl A. EyeSeeThrough: unifying tool selection and application in virtual environments. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). Osaka, Japan, IEEE, 2019, 474–483  
DOI:10.1109/vr.2019.8797988
- 3 Guenter B, Finch M, Drucker S, Tan D, Snyder J. Foveated 3D graphics. *ACM Transactions on Graphics*, 2012, 31(6): 164  
DOI:10.1145/2366145.2366183
- 4 Patney A, Salvi M, Kim J, Kaplanyan A, Wyman C, Benty N, Luebke D, Lefohn A. Towards foveated rendering for

- gaze-tracked virtual reality. *ACM Transactions on Graphics*, 2016, 35(6): 1–12  
DOI:10.1145/2980179.2980246
- 5 Alghofaili R, Solah M S, Huang H K, Sawahata Y, Pomplun M, Yu L F. Optimizing visual element placement via visual attention analysis. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). Osaka, Japan, IEEE, 2019, 464–473  
DOI:10.1109/vr.2019.8797816
  - 6 Hu Z M, Zhang C Y, Li S, Wang G P, Manocha D. SGaze: a data-driven eye-head coordination model for realtime gaze prediction. *IEEE Transactions on Visualization and Computer Graphics*, 2019, 25(5): 2002–2010  
DOI:10.1109/tvcg.2019.2899187
  - 7 Berton F, Olivier A H, Bruneau J, Hoyet L, Pettre J. Studying gaze behaviour during collision avoidance with a virtual walker: influence of the virtual reality setup. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). Osaka, Japan, IEEE, 2019, 717–725  
DOI:10.1109/vr.2019.8798204
  - 8 Chen J, Mi L T, Chen C P, Liu H W, Jiang J H, Zhang W B. Design of foveated contact lens display for augmented reality. *Optics Express*, 2019, 27(26): 38204–38219  
DOI:10.1364/oe.381200
  - 9 Zhou L, Chen C P, Wu Y S, Zhang Z L, Wang K Y, Yu B, Li Y. See-through near-eye displays enabling vision correction. *Optics Express*, 2017, 25(3): 2130–2142  
DOI:10.1364/oe.25.002130
  - 10 Itti L. Models of bottom-up and top-down visual attention. California Institute of Technology. 2000
  - 11 Connor C E, Egeth H E, Yantis S. Visual attention: bottom-up versus top-down. *Current Biology*, 2004, 14(19): R850–R852  
DOI:10.1016/j.cub.2004.09.041
  - 12 Pinto Y, van der Leij A R, Sligte I G, Lamme V A F, Scholte H S. Bottom-up and top-down attention are independent. *Journal of Vision*, 2013, 13(3): 16  
DOI:10.1167/13.3.16
  - 13 Rottach K G, von Maydell R D, Das V E, Zivotofsky A Z, Discenna A O, Gordon J L, Landis D M D, Leigh R J. Evidence for independent feedback control of horizontal and vertical saccades from Niemann-Pick type C disease. *Vision Research*, 1997, 37(24): 3627–3638  
DOI:10.1016/s0042-6989(96)00066-1
  - 14 Sitzmann V, Serrano A, Pavel A, Agrawala M, Gutierrez D, Masia B, Wetzstein G. Saliency in VR: how do people explore virtual environments? *IEEE Transactions on Visualization and Computer Graphics*, 2018, 24(4): 1633–1642  
DOI:10.1109/tvcg.2018.2793599
  - 15 Henderson J. Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, 2003, 7(11): 498–504  
DOI:10.1016/j.tics.2003.09.006
  - 16 Henderson J M, Nuthmann A, Luke S G. Eye movement control during scene viewing: Immediate effects of scene luminance on fixation durations. *Journal of Experimental Psychology: Human Perception and Performance*, 2013, 39(2): 318–322  
DOI:10.1037/a0031224
  - 17 Henderson J M, Olejarczyk J, Luke S G, Schmidt J. Eye movement control during scene viewing: Immediate degradation and enhancement effects of spatial frequency filtering. *Visual Cognition*, 2014, 22(3/4): 486–502  
DOI:10.1080/13506285.2014.897662
  - 18 Cheng M M, Zhang G X, Mitra N J, Huang X L, Hu S M. Global contrast based salient region detection. In: CVPR 2011. Colorado Springs, CO, USA, IEEE, 2011: 409–416  
DOI:10.1109/cvpr.2011.5995344
  - 19 Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(11): 1254–1259  
DOI:10.1109/34.730558

- 20 Borji A, Sihite D N, Itti L. Probabilistic learning of task-specific visual attention. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, IEEE, 2012, 470–477  
DOI:10.1109/cvpr.2012.6247710
- 21 Harel J, Koch C, Perona P. Graph-based visual saliency. In: Advances in neural information processing systems. 2007, 545–552  
DOI: 10.7551/mitpress/7503.003.0073
- 22 Marcella C, Lorenzo B, Giuseppe S, Rita C. Predicting human eye fixations via an LSTM-based saliency attentive model. IEEE Transactions on Image Processing, 2018, 27(10): 5142–5154  
DOI:10.1109/tip.2018.2851672
- 23 Koulieris G A, Drettakis G, Cunningham D, Mania K. Gaze prediction using machine learning for dynamic stereo manipulation in games. In: 2016 IEEE Virtual Reality (VR). Greenville, SC, USA. IEEE, 2016, 113–120  
DOI:10.1109/vr.2016.7504694
- 24 Arabadzhyska E, Tursun O T, Myszkowski K, Seidel H P, Didyk P. Saccade landing position prediction for gaze-contingent rendering. ACM Transactions on Graphics, 2017, 36(4): 1–12  
DOI:10.1145/3072959.3073642
- 25 Box G E, Jenkins G M, Reinsel G C. Time series analysis: forecasting and control. John Wiley & Sons, 2015
- 26 Lachenbruch P A, Cohen J. Statistical power analysis for the behavioral sciences (2nd Ed.). Journal of the American Statistical Association, 1989, 84(408): 1096  
DOI:10.2307/2290095
- 27 Rumsey D J. Statistics II for dummies. John Wiley & Sons, 2009