• **Article** •

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

Zhiming HU [1], Sheng LI[1,2*], Meng GAI[1,2]

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: **Accepted:**

Supported by the National Key R&D Program of China (No.2017YFB0203000) and National Natural Science Foundation of China (No.61632003, No.61661146002, No.61631001).

Abstract: Eye tracking technology gets more and more attention in the field of virtual reality. Specifically, future gaze prediction is crucial in pre-computation for many applications like gaze-contingent rendering, advertisement placement, content-based design, etc. To explore future gaze prediction, it is necessary to analyze the temporal continuity of visual attention in immersive virtual reality. In this paper, we first present the concept of temporal continuity of visual attention. Then we propose a method, i.e. autocorrelation function, to evaluate the temporal continuity. Next, we analyze the temporal continuity in both free-viewing conditions and task-oriented conditions. Specifically, in free-viewing conditions, we perform analysis of a free-viewing gaze dataset and find that the temporal continuity performs well only within a short time interval. In task-oriented conditions, we create a task-oriented game scene and conduct a user study to collect users' gaze data. We analyze the collected gaze data and find the temporal continuity has similar performance with that in free-viewing conditions. Temporal continuity can be applied to future gaze prediction. If the temporal continuity is good, we can directly utilize users' current gaze positions to predict their gaze positions in the future. We further evaluate current gaze's future prediction performances in both free-viewing conditions and task-oriented conditions and find that 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 it has many important applications in the area of virtual reality (VR) including eye movement-based interaction[1, 2], gaze-contingent rendering[3, 4], gaze behavior analysis[5-7], foveated imaging[8], etc. Eye tracking methods can be classified into realtime gaze prediction methods and future gaze prediction methods. Currently, the most common solution for realtime gaze prediction is based on eye trackers. An eye tracker is a hardware device and it can be integrated with head mounted devices[9] (HMDs). Besides eye trackers, software-based solution is also proposed for realtime gaze prediction in virtual reality[6]. However, compared with realtime gaze prediction, there is limited work on future gaze prediction. Future gaze prediction is crucial in pre-computation for many applications like gaze-contingent rendering, advertisement placement, content-based recommendation, etc. To explore the topic of future gaze prediction, there is a necessity to analyze the temporal continuity of visual attention.

In this paper, we present the concept of temporal continuity of visual attention in immersive virtual reality. The temporal continuity refers to the continuity and consistency of users' on-screen gaze position sequence. We propose to utilize autocorrelation function (ACF) to evaluate the temporal continuity. As revealed in prior works[10, 11], there exist two mechanisms of visual attention: a top-down mechanism and a bottom-up mechanism. The temporal continuity of visual attention under top-down mechanism may have different performance from that under bottom-up mechanism. Therefore, we analyze the temporal continuity under the 2 mechanisms independently. Specifically, we explore free-viewing conditions (bottom-up mechanism) and task-oriented conditions (top-down mechanism), respectively.

Free-Viewing Conditions: In free-viewing conditions, we perform analysis of a free-viewing gaze dataset[6] (Section 4). We calculate the ACF of users' gaze position sequence to evaluate the temporal continuity and find that the ACF only performs well within $100\ ms$. The ACF deteriorates significantly with the increase of time interval and it becomes very small when the time interval is larger than $700\ ms$.

Task-Oriented Conditions: To analyze the temporal continuity of visual attention in task-oriented conditions, we create a task-oriented game scene and conduct a user study to collect 19 players' gaze data (Section 5). We analyze the temporal continuity using the collected gaze data and find that the ACF in task-oriented conditions has similar characteristics with that in free-viewing conditions.

Future Gaze Prediction: We further apply the temporal continuity of visual attention to the task of future gaze prediction. If the temporal continuity is good, we can directly employ users' current gaze positions to predict their gaze positions in the future. We find that, in both free-viewing conditions (Section 4.3) and task-oriented conditions (Section 5.3), current gaze only performs well in the task of short-term gaze prediction and it cannot efficiently handle the situation of long-term gaze prediction.

Overall, our contributions include:

- We present the concept of temporal continuity of visual attention in immersive virtual reality with a method to evaluate it.
- We analyze the temporal continuity of visual attention in both free-viewing conditions and task-oriented conditions.
- We apply the temporal continuity to future gaze prediction and evaluate its performance.

2 Related work

In this section, we give 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: a bottom-up mechanism and a top-down mechanism[10, 11]. The bottom-up mechanism is fast and it 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[12]. In addition, the horizontal and vertical eye movements are found to behave differently[13].

Human visual attention has also been studied in the field of virtual reality and it 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[14]. Hu et al. revealed that there exists a linear correlation between users' gaze positions and their head rotation velocities and they further employed users' head movements to predict their realtime gaze positions[6]. In this paper, we focus 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[15]. 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 low-level features of the scene like luminance[16], which impact bottom-up processing, and high-level features[17], 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"[14]. 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[6]. They reported that saccades, which refer to fast eye movements, seldom occur in free-viewing conditions. In this paper, we focus 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[18, 19], which focus on low-level image features like intensity, color, and orientation, or top-down models[20, 21], 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[22].

In the area of virtual reality, however, there is not too much work on gaze prediction. Sitzmann et al. focused on the saliency in 360° static images[14]. They conducted a user 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[23]. They proposed a machine learning-based method to predict the object categories that users gazed at during game play. Hu et al. concentrated on realtime gaze prediction in virtual reality in free-viewing conditions[6]. They proposed an eye-head coordination model for realtime gaze prediction. In this paper, we explore the feasibility of future gaze prediction in immersive virtual reality.

3 Temporal continuity of visual attention

In this section, we first clarify the concept of temporal continuity, including temporal continuity's definition, importance, and application. Then we present a method to evaluate the temporal continuity.

3.1 The concept of temporal continuity

In our research, we define the temporal continuity of visual attention in immersive virtual reality as the continuity and consistency of users' on-screen gaze position sequence. We employ a Cartesian coordinate to describe users' on-screen gaze data by setting the origin to the center of the HMD's screen and orienting the X-axis from left to right and the Y-axis from bottom to top. As utilized in prior works[6, 24], we measure users' horizontal and vertical on-screen gaze position using visual angle, i.e. the angle between user's line of sight and the normal direction of the HMD's screen plane. For example, if a user fixates on the HMD's screen center, his on-screen gaze position will be (0°, 0°).

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, we can directly employ users' current gaze positions to predict their gaze positions in the future:

$$x_g(t_0 + \Delta t) = x_g(t_0),$$
$$y_g(t_0 + \Delta t) = y_g(t_0), \quad (1)$$

where $x_g(t_0)$ and $y_g(t_0)$ are the current horizontal and vertical gaze positions; t_0 is the current time; $x_g(t_0 + \Delta t)$ and $y_g(t_0 + \Delta t)$ are the gaze positions in the future; Δt is the time interval. One of our goals is to determine the value of Δt, at 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 sequence has autocorrelation. Therefore, to evaluate the temporal continuity, we calculate the autocorrelation function (ACF) of users' gaze position sequence by estimating the correlation between gaze position sequence and a delayed copy of the sequence. Specifically, we calculate the ACF of users' horizontal and vertical gaze position sequence, respectively, using the formula proposed in Box et al.' work[25]:

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

In Equation 2, 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 of r_k of 0.1 is classified as small, an absolute value of 0.3 is classified as medium, of 0.5 is classified as strong, and of 0.7 is classified as high[26, 27]. 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; $\bar{y}$ 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 our calculation, we set $k = 0, 1, 2, \dots, 100$ to calculate the autocorrelation. Since the gaze position sequence analyzed in our research is sampled at every $10\,ms$, the time intervals between y_t and y_{t+k} for $k = 0, 1, 2, \dots, 100$ are $0\,ms, 10\,ms, 20\,ms, \dots, 1000\,ms$.

4 Free-viewing conditions

In this section, we analyze the temporal continuity of visual attention in free-viewing conditions. Specifically, we perform analysis of a free-viewing gaze dataset[6]. We first calculate the autocorrelation function of the gaze position sequence to assess the temporal continuity. Then we apply the temporal continuity to future gaze prediction and evaluate its performance.

4.1 Gaze data

Hu et al. recently studied human gaze behaviors in free-viewing conditions in immersive virtual reality and they built a large eye tracking dataset[6], which contains 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 4,000,000 gaze positions and it is large enough to be employed for gaze behavior analysis. Therefore, for simplicity, we directly analyze the temporal continuity based on this dataset.

4.2 Temporal continuity evaluation

To evaluate the temporal continuity, we employ Equation 2 to calculate the autocorrelation function of users' gaze position sequence. Since users' horizontal and vertical gaze behaviors are different[13], we estimate the ACFs of the horizontal and vertical gaze position sequence, respectively. Since there are 240 pieces of data in total, we first calculate the ACF of each piece of data. Then we calculate the mean of the 240 ACFs and utilize the mean ACF as the ACF of users' gaze position sequence in free-viewing conditions. Figure 1 illustrates the horizontal and vertical autocorrelation functions. We can see that both the horizontal and vertical ACFs decrease with the increase of lag time. Within the range of $100\ ms$, the horizontal and vertical ACFs are very high. The horizontal ACF is larger than 0.75 and the vertical ACF is larger than 0.7. When the lag increases to $400\ ms$, 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 $700\ ms$, 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.

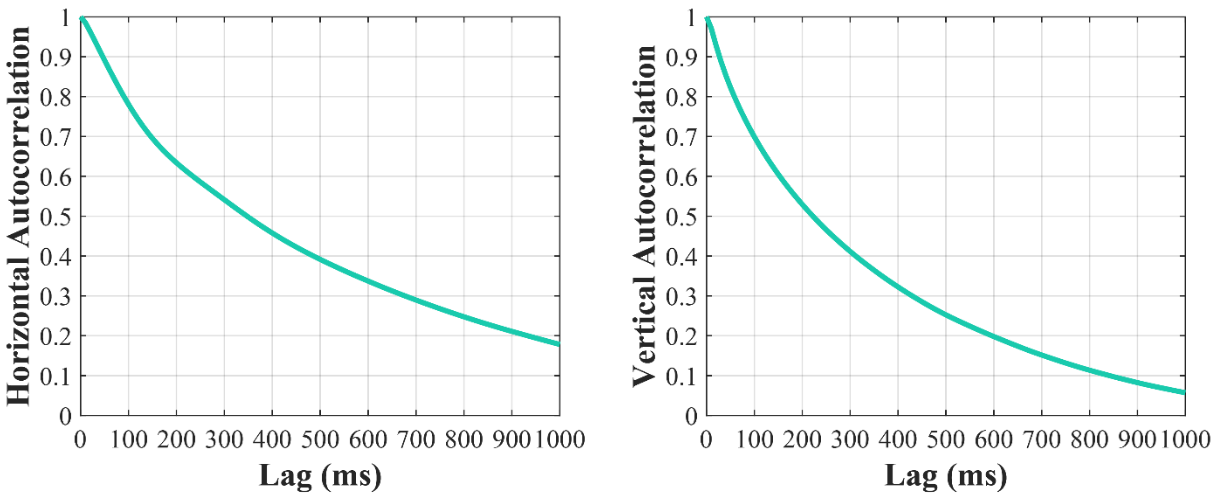

Figure 1 The horizontal (left) and vertical (right) autocorrelation functions of users' gaze position sequence in free-viewing conditions. The horizontal and vertical ACFs decrease when the lag increases. In both the horizontal and vertical directions, the ACFs remain very high within the range of $100\ ms$; the ACFs decrease dramatically when the lag increases to $400\ ms$; the values of the ACFs become very small when the lag is larger than $700\ ms$.

The above analysis reveals the characteristics of the temporal continuity of visual attention in free-viewing conditions. We can conclude that the temporal continuity performs well within a short time ($100\ ms$ or less);

the continuity decreases significantly when the time interval increases; the continuity becomes very weak after a long time ($700\ ms$ 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, we can directly utilize users' current gaze positions to predict their gaze positions in the future (Equation 1). To better evaluate the performance of gaze position prediction, we set an evaluation metric. Specifically, we utilize the **angular distance** between the ground truth and the predicted gaze position, i.e. the angle between user's ground truth line of sight and the predicted line of sight, as the evaluation metric. The smaller the angular distance, the smaller the prediction error and the better the performance. In addition, we also utilize the two baselines that were proposed in Hu et al.'s work[6] as our baselines. Specifically, we utilize the screen center (**Center Baseline**), which is $(0°,\ 0°)$, and the mean of all the gaze positions (**Mean Baseline**), which is $(0.13°, -2.32°)$, as our baselines.

To evaluate temporal continuity's performance on future gaze prediction, we employ current gaze positions, Center baseline, and Mean baseline to predict gaze positions in the future $50\ ms$, $100\ ms$, $150\ ms$, …, $1000\ ms$ and calculate their mean prediction errors (mean angular distances). Figure 2 illustrates the prediction results. We find that Center baseline and Mean baseline retain the same performance at different prediction times because they are constant. Current gaze shows good performance within $100\ ms$. Its prediction performance deteriorates significantly with the increase of prediction time. At the prediction time of $600\ ms$, it even performs 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 the situation of long-term gaze prediction. The left of Figure 3 illustrates a user's gaze trajectory in free-viewing conditions.

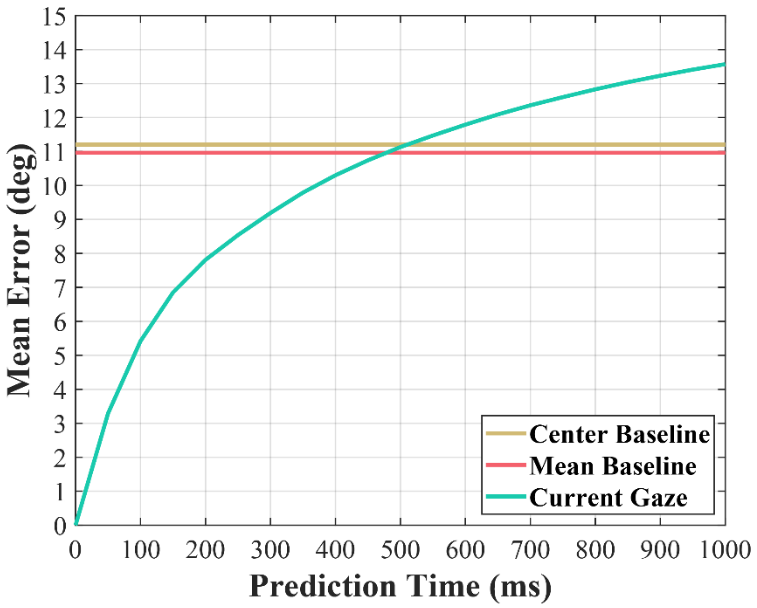

Figure 2 Current gaze and the baselines' future prediction performances in free-viewing conditions. Center baseline and Mean baseline have invariant prediction performances at different prediction times; current gaze shows good performance in the situation of short-term gaze prediction but it does not perform well in the task of long-term prediction.

Figure 3 Users' gaze trajectories in free-viewing conditions (left) and task-oriented conditions (right). The green cross refers to users' current gaze positions; the blue cross indicates users' gaze positions in future 100 ms; the yellow cross denotes gaze positions in future 400 ms; the red cross represents gaze positions in future 700 ms. Current gaze is efficient in short-term gaze prediction but it cannot handle the situation of long-term gaze prediction.

5 Task-oriented conditions

In this section, we analyze the temporal continuity of visual attention in task-oriented conditions. Human visual attention in task-oriented conditions is different from that in free-viewing conditions in that users' 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, we first create a task-oriented game scene to collect users' gaze data. We further perform autocorrelation analysis of our data to evaluate the temporal continuity and also measure temporal continuity's future gaze prediction performance.

5.1 Gaze data

To explore task-oriented conditions, we create a task-oriented game and conduct a user study to collect users' gaze data.

Stimuli: we create a game scene using the Unity game engine and randomly place some animals (ibexes, deer, etc.) in it. The animals are dynamic and their movements are controlled by their own animations. We control the animals' paths using our own Unity script, 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 our game is demonstrated in the left of Figure 4.

Figure 4 Left: Our task-oriented game scene. We randomly place some animals (ibexes, deer, etc.) in the scene and utilize these animals as the targets in the game. Players are given a task to hit these animals with a wand and they will get higher scores if they hit more animals. Right: Our experimental setup.

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

System Details: In our user study, we employ an HTC Vive as our HMD to display our game and also utilize a Vive controller for user interaction. We utilize a 7invensun VR eye tracker, which has a sampling frequency of $100\ Hz$ and has an accuracy of $0.5°$, to collect users' gaze data. Our game scene is created using the Unity game engine. The CPU and GPU of our platform are an Intel(R) Core(TM) i7-8700 @ 3.20GHz and an NVIDIA GeForce RTX 2080 Ti, respectively. The snapshot of our experimental setup is demonstrated in the right of Figure 4.

Procedure: The players were given a Vive controller to help teleport themselves in 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 get familiar with our experimental system. The players were asked to engage in the game for at least 2 minutes and, during the game, we collected their gaze data for later analysis.

Gaze Data: 19 players participated in our game and thus we have 19 pieces of data in total. Each piece of data contains at least 12,000 gaze positions and we have a total of around 300,000 gaze positions.

5.2 Temporal continuity evaluation

To evaluate the temporal continuity, we perform autocorrelation analysis of the gaze data collected in our user study by utilizing Equation 2. We first calculate the ACFs of the 19 pieces of data and then calculate the mean ACF. We utilize the mean ACF as the ACF of users' gaze position sequence 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 $100\ ms$; the ACFs deteriorate significantly with the increase of lag; when the lag is larger than $700\ ms$, the values of the ACFs become very small (horizontal ACF < 0.3, vertical ACF < 0.25).

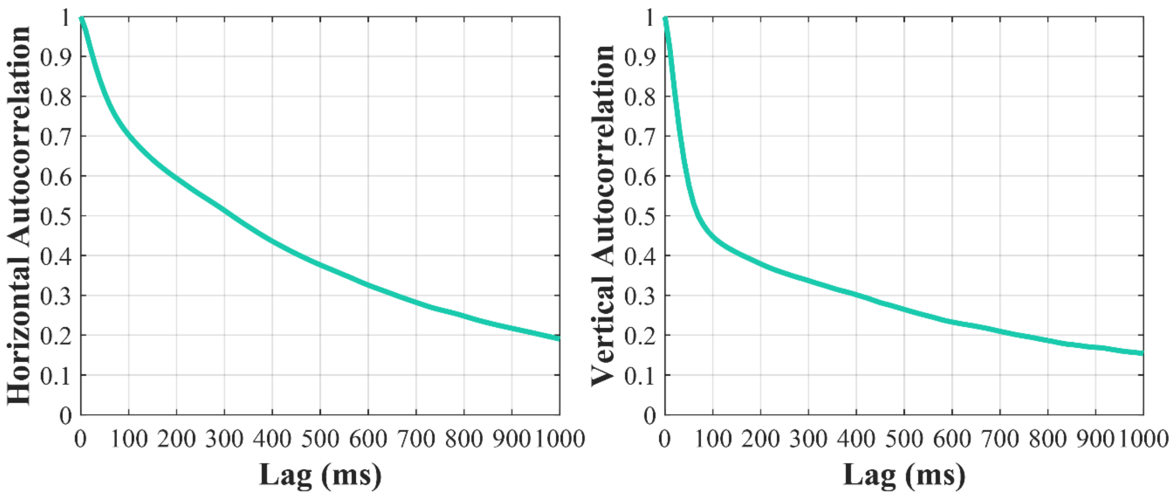

Figure 5 The horizontal (left) and vertical (right) autocorrelation functions of users' gaze position sequence in task-oriented conditions. Both the horizontal and vertical ACFs decrease with the increase of lag. When the lag is small ($\leq 100\ ms$), the ACFs have relatively high values. When the lag becomes very large ($\geq 700\ ms$), the values of the ACFs become very small.

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

5.3 Future gaze prediction

We also evaluate the future gaze prediction performance of current gaze positions based on the gaze data collected in our game. We utilize angular distance as the evaluation metric and employ Center baseline and Mean baseline as our baselines. In this case, Mean baseline refers to the mean of all our gaze data, which is $(4.88°,\ 0.97°)$. We calculate current gaze, Center baseline, and Mean baseline's gaze prediction performances in the future $50\ ms$, $100\ ms$, $150\ ms$, …, $1000\ ms$. Figure 6 illustrates the mean prediction errors of current gaze and the baselines. We can see that the performances of Center baseline and Mean baseline are constant. Within $100\ ms$, current gaze retains high accuracy. However, with the increase of prediction time, the accuracy of current gaze deteriorates significantly. Current gaze preforms worse than Mean baseline when the prediction time is larger than $700\ ms$. These results indicate that the temporal continuity is only effective in the situation of short-term gaze prediction (100 ms or less). A user's gaze trajectory in task-oriented conditions is illustrated in the right of Figure 3.

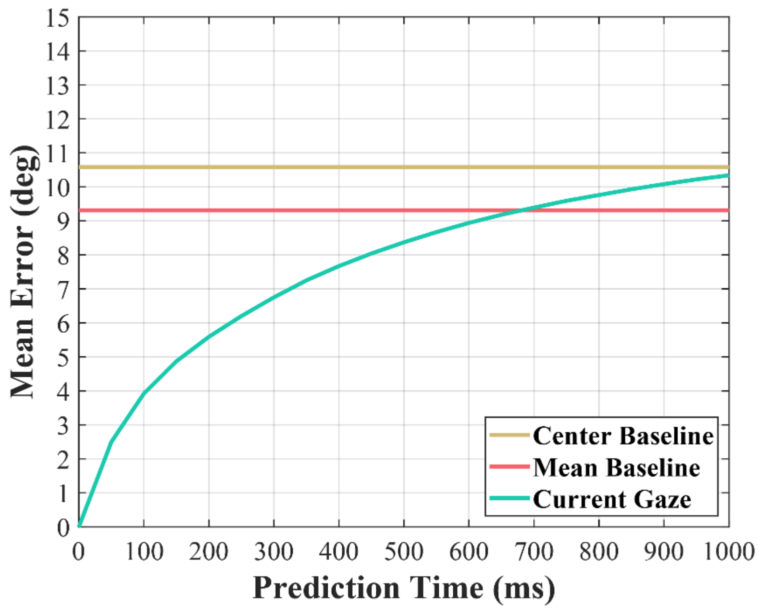

Figure 6 Current gaze and the baselines' future prediction performances in task-oriented conditions. The performances of the baselines are invariant; current gaze shows good performance within a short prediction time and its performance deteriorates dramatically when the prediction time increases.

6 Conclusion, limitations, and future work

In this paper, we present the concept of temporal continuity of visual attention in immersive virtual reality and conduct novel analysis of the temporal continuity. We evaluate the temporal continuity in both free-viewing conditions and task-oriented conditions by calculating autocorrelation functions of users' gaze position sequence. In free-viewing conditions, we perform autocorrelation analysis of a free-viewing gaze dataset and find that the autocorrelation performs well only when the lag is small. In task-oriented conditions, we create a game scene and conduct a user study to collect users' gaze data. We then calculate the autocorrelation functions based on the collected gaze data and find that the temporal continuity in task-oriented conditions is similar to that in free-viewing conditions. We further apply the temporal continuity to the task of future gaze prediction, i.e. utilize current gaze positions to predict gaze positions in the future. Future gaze prediction is vital in pre-computation for many applications like gaze-contingent rendering, advertisement placement, content-based recommendation, etc. The gaze prediction performances reveal that, in both free-viewing conditions and task-oriented conditions, the temporal continuity can only efficiently facilitate short-term gaze prediction. With the increase of prediction time, the efficiency of the temporal continuity deteriorates significantly. The task of long-term gaze prediction still remains to be explored.

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

References

1 Duchowski A T. Gaze-based interaction: A 30 year retrospective. Computers & Graphics-Uk, 2018, 73: 59-69
DOI: 10.1016/j.cag.2018.04.002

2 Mardanbegi D, Mayer B, Pfeuffer K, et al. EyeSeeThrough: Unifying Tool Selection and Application in Virtual Environments. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). 2019. p. 474-483
DOI: 10.1109/vr.2019.8797988

3 Guenter B, Finch M, Drucker S, et al. Foveated 3D Graphics. Acm Transactions on Graphics, 2012, 31(6): 164
DOI: 10.1145/2366145.2366183

4 Patney A, Salvi M, Kim J, et al. 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, et al. Optimizing Visual Element Placement via Visual Attention Analysis. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). 2019. p. 464-473
DOI: 10.1109/vr.2019.8797816

6 Hu Z, Zhang C, Li S, et al. 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, et al. 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). 2019. p. 717-725
DOI: 10.1109/vr.2019.8798204

8 Chen J, Mi L, Chen C P, et al. 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, et al. See-through near-eye displays enabling vision correction. Opt Express, 2017, 25(3): 2130-2142
DOI: 10.1364/OE.25.002130

10 Itti L, Models of bottom-up and top-down visual attention. 2000, California Institute of Technology.

11 Connor C E, Egeth H E, Yantis S. Visual attention: bottom-up versus top-down. Current Biology, 2004, 14(19): R850-852
DOI: 10.1016/j.cub.2004.09.041

12 Pinto Y, van der Leij A R, Sligte I G, et al. 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, et al. Evidence for independent feedback control of horizontal and vertical saccades from Niemann-Pick type C disease. Vision Res, 1997, 37(24): 3627-3638
DOI: 10.1016/S0042-6989(96)00066-1
14 Sitzmann V, Serrano A, Pavel A, et al. 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 M. 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. J Exp Psychol Hum Percept Perform, 2013, 39(2): 318-322
DOI: 10.1037/a0031224
17 Henderson J M, Olejarczyk J, Luke S G, et al. 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, et al. Global contrast based salient region detection. In: Computer Vision and Pattern Recognition (CVPR). 2011. p. 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: Computer Vision and Pattern Recognition (CVPR). 2012, IEEE. p. 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. p. 545-552
DOI: 10.7551/mitpress/7503.003.0073
22 Cornia M, Baraldi L, Serra G, et al. Predicting Human Eye Fixations via an LSTM-based Saliency Attentive Model. IEEE Trans Image Process, 2018, 27(10): 5142 - 5154
DOI: 10.1109/TIP.2018.2851672
23 Koulieris G A, Drettakis G, Cunningham D, et al. Gaze prediction using machine learning for dynamic stereo manipulation in games. In: Virtual Reality (VR). 2016, IEEE. p. 113-120
DOI: 10.1109/vr.2016.7504694
24 Arabadzhiyska E, Tursun O T, Myszkowski K, et al. 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, et al. Time series analysis: forecasting and control. John Wiley & Sons, 2015.
26 Cohen J. Statistical power analysis for the behavioral sciences. 2nd. Hillsdale, NJ: erlbaum, 1988.
27 Rumsey D J. Statistics II for dummies. John Wiley & Sons, 2009.