

Advanced Gaze Visualizations for Three-dimensional Virtual Environments

Sophie Stellmach*, Lennart Nacke**, Raimund Dachzelt*

*User Interface & Software Engineering Group
Faculty of Computer Science
Otto-von-Guericke-Universität Magdeburg
D-39106 Magdeburg, Germany
stellmach@acm.org, dachzelt@acm.org

**Game Systems and Interaction Research Laboratory
School of Computing
Blekinge Institute of Technology
SE-37225 Ronneby, Sweden
lennart.nacke@acm.org

Abstract

Gaze visualizations represent an effective way for gaining fast insights into eye tracking data. Current approaches do not adequately support eye tracking studies for three-dimensional (3D) virtual environments. Hence, we propose a set of advanced gaze visualization techniques for supporting gaze behavior analysis in such environments. Similar to commonly used gaze visualizations for two-dimensional stimuli (e.g., images and websites), we contribute *advanced 3D scan paths* and *3D attentional maps*. In addition, we introduce a *models of interest timeline* depicting viewed models, which can be used for displaying scan paths in a selected time segment. A prototype toolkit is also discussed which combines an implementation of our proposed techniques. Their potential for facilitating eye tracking studies in virtual environments was supported by a user study among eye tracking and visualization experts.

CR Categories: H.5.1 [Multimedia Information Systems]: Evaluation/Methodology—Information Interfaces and Presentation I.4.8 [Computing Methodologies]: Image Processing and Computer Vision—Scene Analysis

Keywords: Gaze visualizations, eye tracking, eye movements, attentional maps, scan paths, virtual environments, three-dimensional

1 Introduction

For investigating gaze behavior, large sets of numerical data have to be collected and evaluated, such as eye positions, calculated fixations and saccades, as well as total fixation times [Ramloll et al. 2004]. Consequently, gaze visualizations have helped detecting, understanding, and interpreting essential relationships in such multivariate and multidimensional data. Especially superimposed techniques proved useful for imparting how a viewer’s gaze behaved with respect to an underlying stimulus. Common two-dimensional (2D) gaze visualization techniques include scan paths, attentional maps, and areas of interest.

While traditional diagnostic eye tracking studies have focused on a person’s gaze behavior while reading texts or observing images, new application areas have emerged in connection with the advancement of technology. This includes the perception of digital content such as websites, user interfaces, digital games, and three-dimensional (3D) virtual environments (VE). However, current gaze visualization techniques do not adequately support interactive visual analysis of 3D VEs. On the one hand, they are not well

adapted to challenges implied by 3D space (e.g., occlusion problems or varying viewpoints). On the other hand, they do not allow for quickly determining which parts of a 3D model were viewed the longest. After all, this lack of suitable techniques currently results in a time-consuming evaluation of eye tracking studies in 3D VEs. Usually, video recordings of displayed content have to be examined frame-by-frame, for which gaze data are represented as superimposed gaze plots. Thus, more suitable techniques are required for a faster visual gaze analysis in various application domains (e.g., virtual interactive trainings [Duchowski et al. 2000], social networking environments, and computer-aided design). Such techniques could, for example, assist in quickly examining 3D product design features or people’s attentional foci when being confronted with a critical situation.

This paper makes a methodological and technological contribution to the investigation of visual attention in 3D VEs. To facilitate eye tracking studies in 3D VEs, we propose a set of adapted and novel gaze visualization techniques: superimposed *3D scan paths* and *3D attentional maps*, as well as a *models of interest timeline* depicting viewed models. We have also developed a gaze analysis tool for the evaluation of eye tracking studies in static 3D VEs. The techniques and the tool are described together with related work in the following section. In Section 3, a survey is discussed which was conducted among eye tracking and visualization experts for assessing the usefulness of our proposed techniques.

2 Gaze Visualization Techniques for Three-dimensional Virtual Environments

Common gaze visualization techniques for evaluating eye tracking studies include scan paths and attentional maps but fall short when it comes to 3D VEs. For example, the gaze behavior of customers in a virtual 3D store could be evaluated with regard to product and advertisement placements. Potential research questions could ask in which order items were observed or how visual interest is generally distributed. For answering questions like these, we introduce a set of gaze visualization techniques adapted for static 3D VEs. These techniques include superimposed visualizations of *3D scan paths* and *3D attentional maps*, as well as a *models of interest timeline* depicting the order of fixated objects. For them, a 3D scene is assumed with several static models, which users can freely explore and for which sample gaze data was collected. Furthermore, we presume a lack of transparent phenomena (e.g., fog or glass). Our approach uses 3D gaze positions which are calculated as intersections between gaze rays and viewed models (see Section 2.4).

2.1 Three-dimensional Scan Paths

A common approach for depicting eye tracking data are scan path visualizations superimposed on a stimulus representation. Scan paths show the order of eye movements by drawing connected lines (*saccades*) between subsequent gaze positions (*fixations*). *Fixation and saccade plots* imply data clustering by grouping infor-

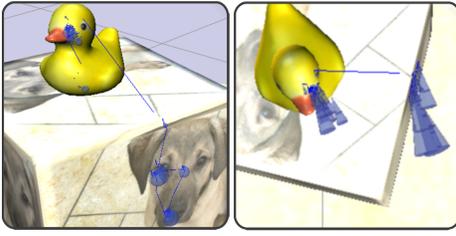


Figure 1: Two alternative fixation representations, spheres (a) and cones (b), are presented for 3D scan paths.

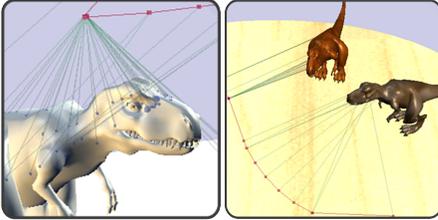


Figure 2: Two examples of camera paths (viewpoints and viewing directions) and fixations.

mation into meaningful units (fixations). While fixations indicate aspects attracting an observer’s attention, saccades provide information about how fixations are related to each other. Thereby, fixations are usually displayed as circles varying in size depending on the fixation duration and saccades as straight lines.

Scan paths are frequently used for static (e.g., images and texts) and dynamic (e.g., videos) 2D stimuli. For other stimuli, superimposed fixation plots are usually applied to recorded videos of the presented content. This often implies a time-consuming frame-by-frame video data analysis. One solution proposed by Ramloll et al. [2004] is a fixation net for dynamic 3D non-stereoscopic content, for which gaze positions are mapped onto flattened objects. Representing binocular scan paths in 3D VRs has been presented by Duchowski et al. [2000] for which 2D gaze positions and depths (gaze depth paths) are depicted.

In contrast to Duchowski et al., we propose a monocular scan path depicting 3D gaze positions as intersections of a gaze ray and a viewed model. Besides traditional spherical representations (see Figure 1a), we used conical fixation representations pointing at a corresponding gaze position (see Figure 1b), since it may integrate additional information about varying camera positions. So, cones could represent gaze positions (apex), fixation durations (cone’s base), viewing directions (cone’s orientation), and viewing distances (cone’s height) within one representation. The traditional saccade representation can cause problems in 3D, because saccade lines may cross through surfaces. A simple solution we propose is to maintain the traditional saccade representation with the possibility for adapting rendering options of 3D models (e.g., wireframe models or hiding objects) for determining linked fixations.

It is also important to visualize how the locations and viewing directions of the virtual camera have changed during observation. This may aid in finding out, for example, if a scene was observed from diverse locations. We propose to visualize the camera path with traces pointing at the respective gaze positions. Figure 2 shows an example for a camera path for which the camera locations are depicted as red lines and viewing directions as straight blue lines. Displayed scan and camera paths can be filtered by means of the *models of interest timeline* to provide a better overview.

2.2 Models of Interest Timeline

A timeline visualization maps data against time and facilitates finding events before, after, or during a given time interval. Thus, it can be used to answer several questions, such as: *Has object x been observed repetitively? In what order and for how long are objects looked at?* In this regard, the *Object vs. Time View* [Lessing and Linge 2002] is a series of timelines for depicting Areas of Interest (AOIs). Each AOI is assigned a time row, resulting in an AOI matrix. With an increasing number of AOIs and shorter fixation times, the table size increases and may hinder a good overview.

Therefore, we propose a space-filling *models of interest (MOI) timeline* (see Figure 3) for a compact illustration that effectively utilizes assigned screen space by omitting voids. We suggest to use the term *Model of Interest* to describe distinct 3D objects. The MOI timeline gives an overview about a user’s gaze distribution based on viewed models. Each model is labeled with a specific color, which can be manually adapted. By assigning the same value to different objects, semantic groups can be defined, for example, for grouping similar looking models, object hierarchies or closely arranged items. A legend is displayed for providing an overview of assigned colors (see in the left area of Figure 5).

Zooming and selection techniques may aid in making data visible which would otherwise be suppressed due to limited display sizes. Based on TimeZoom [Dachselt and Weiland 2006], the MOI timeline can be dragged to a desired position to offer horizontal scrolling and supports continuous zooming (e.g., via a zoom slider, see Figure 3). For this purpose, the viewing area is divided into a large detail area for zooming (see Figure 3) and a small overview area displaying the entire collection. Displayed scan and camera paths can be filtered with respect to a selected period. For this purpose, navigation markers will appear if the path visualizations are employed (see Figure 3), otherwise the markers are hidden. The user can simply click and drag the markers to define intervals of interest within the timeline. While dragging the timeline, the user can play back data of interest as defined by this fixed interval. Finally, additional details for a timeline entry (e.g., object identifiers, start and finish times) can be displayed, when hovering over it with the mouse cursor. Context menus are triggered by right-clicking on items, for instance, for changing its color.

2.3 Three-dimensional Attentional Maps

An attentional map (or heatmap) is an aggregated representation depicting areas of visual attention for primarily static 2D stimuli. This is substantiated by the fact that an attentional map usually has the same dimensions (width and height) as the underlying stimulus [Wooding 2002]. Since fixations are merely accumulated, attentional maps do not provide information about gaze sequences, but are suitable for indicating areas attracting visual attention over a certain period of time. Visualizing gaze data directly in the 3D VE enables an aggregated representation of longer observations (in contrast to the traditional frame-by-frame evaluation).

We propose *3D attentional maps* as superimposed aggregated visualizations for virtual 3D stimuli: *projected*, *object-based*, and *surface-based* attentional representations (see Figure 4). A *projected attentional map* is a 2D representation of 3D gaze distributions for selected arbitrary views (e.g. a top view as in Figure 4). For an *object-based attentional representation* one color is mapped to the surface of each model based on its received visual attention (see Figure 4b). This allows for quickly evaluating an object’s visual attractiveness while maintaining information about the spatial relationship with other models (e.g., two adjacent models have received high visual attention). *Surface-based attentional maps* dis-



Figure 3: An example of the models of interest timeline with selection markers for confining displayed scan and camera paths.

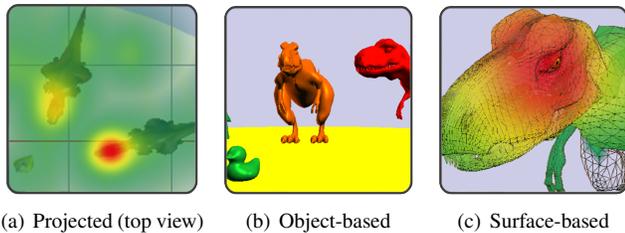


Figure 4: Advanced attentional maps for the application in three-dimensional virtual environments.

play aggregated fixation data as heatmaps directly on a model’s surface using a vertex-based mapping (see Figure 4c). This provides detailed insights into gaze distributions across models’ surfaces, allowing to draw quick conclusions about which model aspects attracted visual interest.

A combination of these techniques offers different levels of detail for data investigation. While a projected heatmap may give an overview of the gaze distribution across a scene, an object-based attentional map can provide information about the models’ visual attractiveness when zooming in. Finally, the surface-based technique allows for close examinations of viewed models.

2.4 Implementation

For the implementation of the presented gaze visualization techniques, a virtual 3D scene was required for which we used Microsoft’s XNA Game Studio 3.0 (based on C#). Our system was confined to static 3D VEs without any transparent phenomena (e.g., smoke or semi-transparent textures). Users can freely explore the scene by moving their camera viewpoints via mouse and keyboard controls. In addition, an integration of the Tobii 1750 Eye Tracker and XNA allowed for logging 3D gaze target positions on models’ surfaces. For this purpose, a 3D collision ray needed to be determined based on 2D screen-based gaze positions which are supplied by the eye tracker. The gaze ray was used to calculate and log its intersection with virtual objects on the precision level of a polygonal triangle on a model [Möller and Trumbore 2005]. The processed data were stored in log files for post-analysis.

The presented visualization techniques were implemented in a prototype of a gaze analysis software tool: SVEETER (see Figure 5). It is based on XNA (for 3D scenes) and Windows Forms to benefit from existing interface elements (e.g., buttons and menus). SVEETER offers a coherent framework for loading 3D scenes and corresponding gaze data logs, as well as deploying adapted gaze visualizations techniques. Multiple views are integrated to look at a scene from different viewpoints. Context menus offer the possibility to apply certain options to each view, such as adapted rendering options. The figures used throughout this paper to illustrate the different gaze visualization techniques were created with SVEETER.

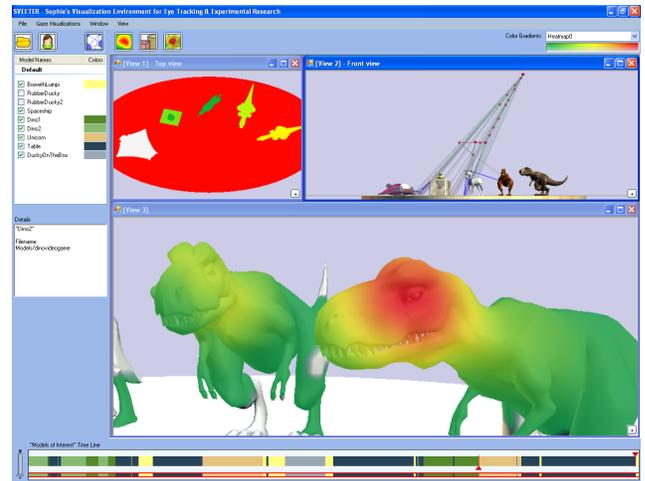


Figure 5: A screenshot from SVEETER illustrating multiple views at a scene. The MOI timeline is shown in the lower area with its legend displayed to the upper left.

3 User Study

We have conducted a user study with eye tracking and visualization experts to assess the usefulness of the presented techniques. The method and results are discussed in the following paragraphs.

Participants. Group 1 consisted of 20 eye tracking professionals and researchers, aged between 23 and 52 years (Mean [M] = 34.50). Participants in this group rated their general eye tracking knowledge¹ higher than average (M = 3.85, Standard Deviation [SD] = 0.96). Group 2 included 8 local data visualization and computer graphics experts, aged between 25 and 35 years (M = 28.25). This group rated their eye tracking knowledge under average (M = 1.25, SD = 1.09), but their expertise in computational visualization above average (M = 3.63, SD = 1.22).

Measures. The usefulness of the gaze visualization techniques was investigated in an online survey². Each technique was briefly described with screenshots. Respondents were asked to rate their agreement to statements such as “Cone visualizations are useful for representing fixations in virtual environments.” The qualitative part of the survey collected comments about usefulness, improvements, and possible applications of the techniques.

Procedure and Design. Group 1 had to answer the questions based on brief textual and pictorial descriptions of the gaze visualization techniques provided in the online survey. For a better understanding of the new gaze visualizations, group 2 could use the SVEETER tool at our University. After welcoming each partici-

¹Rated on a Likert-scale from 1 (poor) to 5 (excellent).

²The survey included 19 Likert scales, from 1 (do not agree at all) to 5 (extremely agree), and 6 qualitative open questions using LimeSurvey (Version 1.80+)

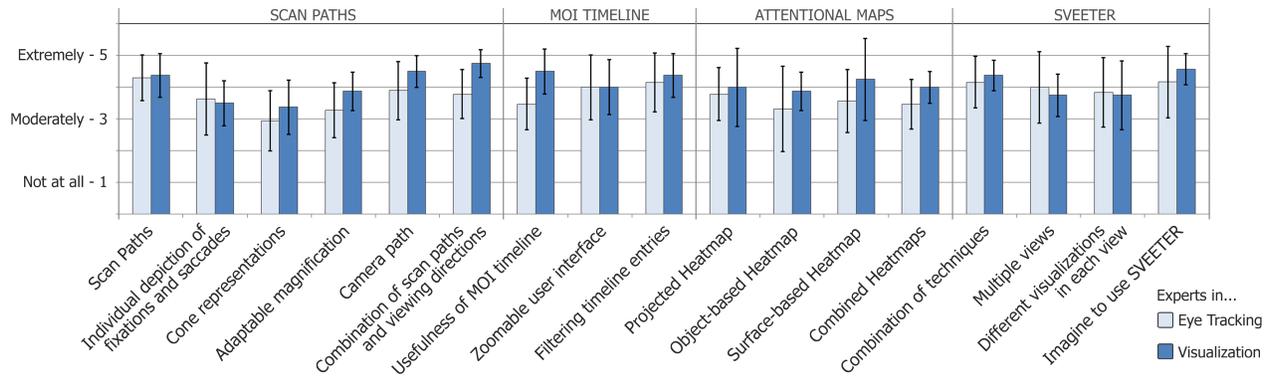


Figure 6: Agreement ratings from both groups with the corresponding standard deviations.

pant of group 2, a short introduction about visual gaze analysis was provided and the tool was briefly presented. Following this, sample gaze data for a 3D scene and a set of 9 predefined tasks were given to each individual. Thereby, the aim was systematical acquaintance with the techniques, not an efficiency evaluation. Tasks included, for example, to find out if particular parts of an object received high visual attention and if observers changed their viewpoints. After completing the tasks and familiarizing themselves with the different techniques, the visualization experts were asked to fill out the online survey. On average, each session took about 45 minutes.

Results. Means and standard deviation values are omitted in this paper, since none of the subjective results showed significant differences between group 1 and group 2. In general, participants agreed that gaze visualization techniques facilitate eye tracking analysis. The detailed results are shown in Figure 6.

Scan paths are generally useful for studying gaze sequences. In this context, depicting camera paths is regarded important as well. Thus, a combination of camera and scan paths is found useful by both groups. In contrast, cones were considered only moderately suitable for representing fixations. Instead, a simple combination of the spherical fixation representations with viewing directions and positions was preferred. While the motivation for individually depicting fixations and saccades was not evident to everybody, it was regarded useful in the overall rating.

Besides limiting scan and camera paths temporally via the MOI timeline, additional filtering functionality was requested. Participants agreed that the MOI timeline helps to detect gaze patterns. Group 2 showed great interest in the MOI timeline when testing SVEETER. Both groups agreed that a zoomable user interface is convenient for the timeline. Filtering objects of interest in the timeline was rated beneficial. Suggested improvements for the timeline included substituting color identification with iconic images and improved suitability for scenes with many objects by grouping them dynamically or individually.

The SVEETER tool was considered highly useful, combining scan paths, attentional maps, and the MOI timeline. Thus, for evaluating eye tracking studies in 3D VEs, interviewees could imagine to use SVEETER. Multiple views are rated practical together with the ability to assign different visualizations to each view. We observed that a combination of different techniques was often employed by users testing SVEETER (group 2). Both scan paths and surface-based maps were, for example, used to investigate observation patterns on a model. Participants frequently faded out objects to gain a better overview. In addition, we noticed that the MOI timeline was often used instead of scan paths.

4 Conclusion and Future Work

In this paper, we presented a set of advanced gaze visualization techniques for investigating visual attention in 3D VEs. *Three-dimensional scan paths* assist in examining sequences of fixations and saccades. Thereby, *camera paths* provide valuable information about how a user navigated through a scene and from which locations objects have been observed. The *models of interest timeline* may help answering, for example, whether an object was viewed at a certain point in time or whether any cyclic viewing behavior existed. Finally, three types of *attentional maps* were discussed to examine how visual attention is distributed across a scene (*projected heatmap*), among 3D models (*object-based attentional map*), and across a model’s surface (*surface-based heatmap*). A combination of these techniques was integrated in a toolkit for visual analysis of gaze data (SVEETER). The survey results from eye tracking and visualization experts indicate the potential usefulness for various application areas.

Since visual gaze analysis for 3D VEs is still in an early stage, our techniques may serve as a solid foundation and provide a basis for further development. This includes representing data from multiple users as well as developing and testing alternative techniques.

References

- DACHSELT, R., AND WEILAND, M. 2006. TimeZoom: a flexible detail and context timeline. In *CHI '06: Extended abstracts on Human factors in computing systems*, ACM, 682–687.
- DUCHOWSKI, A. T., SHIVASHANKARAIHAH, V., RAWLS, T., GRAMOPADHYE, A. K., MELLODY, B., AND KANKI, B. 2000. Binocular eye tracking in virtual reality for inspection training. In *ETRA '00*, ACM, 89–96.
- LESSING, S., AND LINGE, L. 2002. *IICap: A new environment for eye tracking data analysis*. Master’s thesis. University of Lund, Sweden.
- MÖLLER, T., AND TRUMBORE, B. 2005. Fast, minimum storage ray/triangle intersection. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Courses*, ACM.
- RAMLOLL, R., TREPAGNIER, C., SEBRECHTS, M., AND BEEDASY, J. 2004. Gaze data visualization tools: opportunities and challenges. *Eighth International Conference on Information Visualisation* (July), 173–180.
- WOODING, D. S. 2002. Fixation maps: quantifying eye-movement traces. In *ETRA '02*, ACM, 31–36.