3D Attentional Maps - Aggregated Gaze Visualizations in Three-Dimensional Virtual Environments

Sophie Stellmach
Faculty of Computer Science
Otto-von-Guericke University
Magdeburg, Germany
stellmach@acm.org

Lennart Nacke
Department of Computer Science
University of Saskatchewan
Saskatoon, Canada
lennart.nacke@acm.org

Raimund Dachselt
Faculty of Computer Science
Otto-von-Guericke University
Magdeburg, Germany
dachselt@acm.org

ABSTRACT
Gaze visualizations hold the potential to facilitate usability studies of interactive systems. However, visual gaze analysis in three-dimensional virtual environments still lacks methods and techniques for aggregating attentional representations. We propose three novel gaze visualizations for the application in such environments: *projected*, *object-based,* and *surface-based* attentional maps. These techniques provide an overview of how visual attention is distributed across a scene, among different models, and across a model's surface. Two user studies conducted among eye tracking and visualization experts approve the high value of these techniques for the fast evaluation of eye tracking studies in virtual environments.

Categories and Subject Descriptors
H.5.1 [**Multimedia Information Systems**]: Evaluation/ methodology - *Information Interfaces and Presentation*

General Terms
Design, Experimentation

Keywords
Gaze visualizations, eye tracking, attentional map, eye movements, virtual environments, three-dimensional scene

1. INTRODUCTION
Diagnostic eye tracking studies concentrate on post-trial (i.e., offline) assessment of observers' gaze behavior by recording their eye movements [1]. For this purpose, gaze visualizations are frequently used to gain quick insights into large gaze data sets [6]. A common technique for investigating visual attention is the aggregation and representation of gaze target positions in a superimposed attentional map (also commonly referred to as *heat map* or *attentional landscape* [7]). Attentional maps are suitable for static two-dimensional (2D) stimuli, since they are generally superimposed over an underlying stimulus and thus their dimensions (width and height) are the same [8]. They have been used in studies on how images [3, 8], websites, and 2D user interfaces are visually perceived by an observer or user. Each pixel in an attentional map is assigned a value for describing its degree of visual attraction over a certain period of time. This value can be represented as height in three-dimensional (3D) fixation maps or as a color in 2D contour maps [8]. Attentional maps can be customized by adapting the heat signature, opacity, and time interval [3].

While attentional maps are suitable for 2D stimuli, their application for 3D stimuli would lead to data loss, since 3D fixation data would have to be represented in a 2D graph. Three-dimensional virtual environments (VEs) are applied in various contexts, such as 3D gaming (e.g., First-Person Shooters), virtual interactive training, computer-aided design, as well as social networking environments (e.g., Second Life). In order to improve the design and evaluation of such environments it could be vital to learn how certain aspects in a scene affect a user's visual attention. This can also aid in understanding a user's mental processes under particular circumstances (e.g., decision making in critical situations). Until now, such eye tracking studies typically comprise a time-consuming frame-by-frame analysis of captured screen recordings with superimposed scan paths. One of the few available gaze visualization techniques for 3D contexts is the representation of fixations and saccades as 3D scan paths [2, 5]. In a nutshell, while the relevance for studying gaze behavior in VEs increases, adapted gaze visualization techniques are limited [6].

In order to improve visual analysis of gaze data in static 3D VEs, we propose a set of novel aggregated gaze visualizations: *projected*, *object-based*, and *surface-based* attentional maps. By providing different levels of detail of the gaze distribution, a combination of these techniques has the potential to considerably facilitate usability studies in VEs. In the following section, the novel gaze visualizations are introduced. Finally, a study is described that investigates how eye tracking and visualization experts assess the utility of the presented gaze visualizations.

2. ADVANCED ATTENTIONAL MAPS
We introduce three novel gaze visualization techniques for superimposing aggregated fixation data over virtual 3D stimuli. *Projected attentional maps* are 2D overview representations of 3D gaze data. *Object-based attentional maps* assign colors as model textures to indicate visual attractiveness of 3D objects. *Surface-based attentional maps* display gaze fixation data as heat maps on 3D model surfaces. For this purpose, we assume a 3D scene in which the models remain unchanged and which users can freely explore (i.e., dynamic camera control, but static objects). The intersection points of gaze rays with observed virtual objects have been collected as 3D gaze data for post-analysis. It is important to note that attentional maps may represent different aspects of fixation data, such as duration, count or frequency. The proposed visualizations can be adapted for the respective research question.

AVI '10, May 25-29, 2010, Rome, Italy

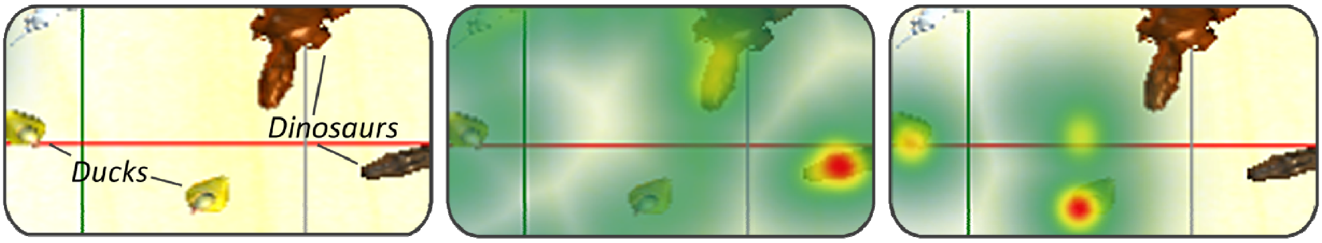

Figure 1. Two examples of a projected attentional map for the top view with different gaze data.

2.1 Projected Attentional Maps

As a direct consequence from 2D eye tracking heat maps, we propose the *projected attentional map*, which is a 2D planar representation of 3D fixation data (see Figure 1) similar to overview maps. It allows a quick overview of the distribution of visual attention across a scene (which the observer could freely explore). However, it is limited in suitability for detailed inspections due to the ambiguous representation, which results from mapping 3D data to a 2D layer.

For the creation of projected attentional maps, 3D fixation positions are required. The occurrences of fixation points are accumulated, which are located on rays orthogonal to a projection plane, which can be defined by the analyst. The projection plane can be defined for standard (e.g., top, front, side) or arbitrary viewpoints. Similar to Wooding's contour plots [8], a 2D Gaussian distribution is assigned. The projected attentional map needs to be recalculated if a new viewpoint is assigned. While its performance is independent from the scene size and number of objects in it, it is accelerated by a decreased amount of gaze data.

2.2 Object-based Attentional Maps

For the *object-based attentional map* we propose representing a model's visual attractiveness by assigning a representative color to its whole textural surface (see Figure 2). This supports rapid detection of objects, which have been looked at, while providing information about how objects are situated towards each other at the same time. Thereby, objects may also be structured within a scene hierarchy. This way, visualizations may not only incorporate single objects but instead semantic groups.

In order to create an object-based attentional map the fixation data from each model are required. A model's color is chosen based on the received visual attention and the selected heat signature. Since only objects' colors have to be adapted and because of the independence from assigned viewpoints the runtime of this visualization technique is very fast. The fixation data, such as the total fixation count, only need to be calculated once and can be stored for reuse. The performance of the initial calculation is affected by the number of objects and the amount of gaze data.

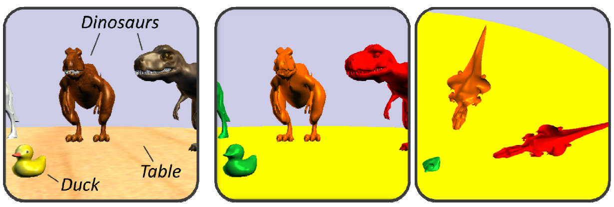

Figure 2. An object-based attentional map for two views.

2.3 Surface-based Attentional Maps

With the *surface-based attentional map* we present a new gaze visualization technique for facilitating detailed inspections of visual attention across 3D models' surfaces (see Figure 3). Aggregated fixation data is depicted in a novel way, i.e., directly on models' textural surfaces that allows for drawing conclusions about which regions of a model attract high visual interest.

For the creation of surface-based attentional maps it has to be determined how often each mesh triangle has been looked at. Since the assembled 3D gaze positions are intersection points of gaze rays with viewed models, each gaze position can be attributed to a particular mesh triangle. A 3D Gaussian distribution is used for spreading gaze information across adjacent mesh triangles. Finally, a triangle mesh is assembled resulting in a "second skin" for each viewed model (i.e., a textural representation of aggregated fixations on models' surfaces).

It has to be noted that the model mesh needs to be carefully chosen for this visualization. If a simple box (as in Figure 3) is described by only the minimal amount of mesh triangles (which is 12) then it is not possible to get a smooth attentional map.

Compared to the previous techniques, the surface-based attentional map is the most complex approach, since models' triangle meshes have to be accessed, which results in longer runtimes. Thus, its performance is affected by the amount of processed gaze data and the complexity of viewed 3D models. However, the calculations only need to be performed when loading new user data, since the technique is independent from changing viewpoints in the analysis software.

2.4 Combined Attentional Maps

Although the individual techniques can be used independently, a combination of the proposed techniques can also be used to implement different levels of detail (LODs) for data examination. A projected attentional map can help identifying areas of high interest if a scene is scrutinized from afar. When zoomed in, distinct objects are colorized based on how often they have been observed (unicolored models). Further amplification shows a detailed representation of how the visual attention is distributed over a model's surface. This may improve performance issues especially with respect to surface-based attentional maps.

2.5 Implementation

The presented techniques have been implemented in a prototype of a gaze analysis software tool for evaluating eye tracking studies in static 3D VEs: SVEETER. It is based on Microsoft's XNA framework and Windows Forms. SVEETER has been used to illustrate the proposed techniques throughout this paper.

Requirements for implementing the presented attentional maps include a virtual scene in 3D space for which gaze target positions are logged on 3D models. Due to the lack of a well-established

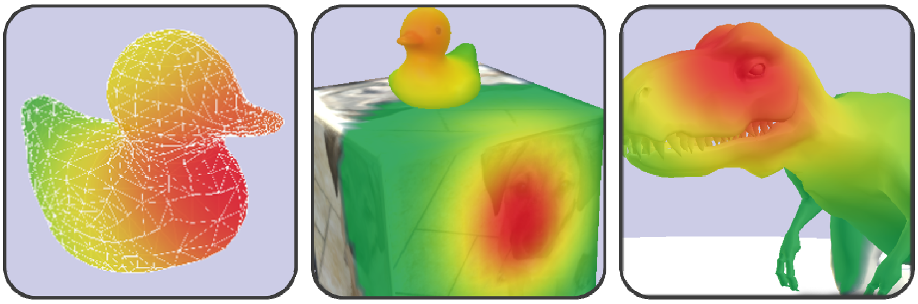

Figure 3. Several surface-based attentional representations with different rendering options.

gaze visualization methodology for representing aggregated fixations in 3D VEs, we have confined our system to static 3D VEs that omit 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.

For collecting 3D gaze target positions, an adapted 3D gaze tracing mechanism was implemented. This algorithm works in the following way: A 3D collision ray needs to be determined based on 2D screen-based gaze positions. 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 [4]. These 3D gaze positions are stored in log files, which can be loaded with SVEETER for deploying the presented visualizations techniques.

3. USER STUDY

In order to find out if our techniques are useful for HCI practitioners and researchers who carry out gaze analysis in VEs and to gather expert feedback, we have conducted a user study with eye tracking and visualization experts. The method and results are reported below.

3.1 Method

3.1.1 Participants

Two groups of participants are distinguished who took part in an online survey about the suitability of the presented techniques for different purposes. While the first group had to answer the questions based on brief descriptions of the techniques provided in the survey, the second group was given the possibility to test these techniques with SVEETER.

The first group consisted of 20 international eye tracking professionals and researchers, aged between 23 and 52 (M = 34.50). On a scale from 1 (poor) to 5 (excellent) participants from this group rated their eye tracking knowledge as higher than average (M = 3.85, SD = 0.96). The second group included 8 experts in the field of data visualization and computer graphics working at our university, aged between 25 and 35 years (M = 28.25). While participants from the second group rated their eye tracking knowledge as lower than average (M = 1.25, SD = 1.09), they assessed their expertise in computational visualization as relatively high (M = 3.63, SD = 1.22).

3.1.2 Measures

The online survey was implemented with the open source survey application LimeSurvey (Version 1.80+). As part of analyzing the SVEETER toolkit, the usefulness of the presented attentional maps was investigated (six rating-scale and two free-text questions). Each technique was briefly described and screenshots were presented. Information about the subjects' background was also gathered (i.e., demographics, eye tracking experience).

Respondents were asked to rate their agreement to statements on a Likert scale from 1 (do not agree at all) to 5 (extremely agree), such as: "*Projected attentional maps provide a good overview of the gaze distribution.*" and "*Object-based attentional maps give a quick overview about objects' visual attractiveness and how they are spatially related to each other.*" The free-text (i.e., qualitative) questions asked for comments about the techniques' usefulness and improvements.

To provide a better understanding of attentional maps, the second group could explore the potential of these techniques by using SVEETER. For testing the proposed visualization techniques, a virtual scene was designed containing several 3D models for which gaze data has been collected with the Tobii 1750 eye tracker. The observer could freely explore the static scene.

3.1.3 Procedure and Design

For the first group, eye tracking experts were invited to participate in the online survey via email based on searches in major eye tracking publication venues, such as COGAIN, ETRA, and ECEM and with the support of staff from Tobii Technology AB. The online survey has been conducted over a time period of 19 days, allowing participants to individually decide, when and how extensive to answer the posed questions.

The second study design (toolkit and online survey – visualization experts group) was conducted locally at our university. After welcoming each participant, a short introduction about visual gaze analysis was provided. After briefly presenting the toolkit, a set of 9 predefined tasks was given to each individual. The tasks aimed at a systematical acquaintance with the featured techniques. However, users could decide individually which visualization seemed most appropriate for answering. The questions focused on how the visual interest was distributed. This included whether particular parts of an object received high visual attention or whether certain areas in a scene were neglected. After completing the tasks and familiarizing themselves with the different techniques, participants from the visualization experts group were asked to fill out the online survey. On average, each session took about 45 minutes with initial preparations, instructions, and the completion of the survey.

3.2 Results

After evaluating the quantitative answers, subjective impressions and suggested improvements are reported.

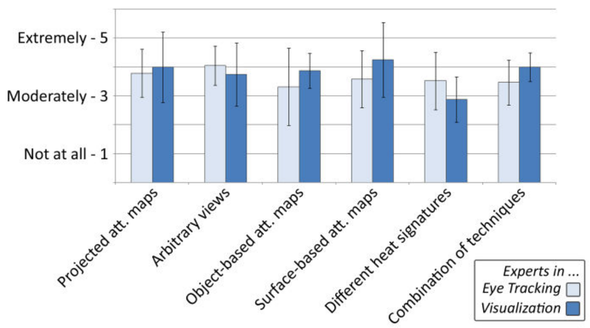

Figure 4. Agreement ratings from both groups with the corresponding standard deviations.

3.2.1 Quantitative Results from Rating Scales

The quantitative results from both groups are illustrated in Figure 4. By conducting several Mann-Whitney tests, we found out that the opinions between both groups did not differ significantly. Thus, the lack of testing the gaze visualizations did not considerably affect the answers.

Participants agreed that *projected attentional maps* provide a good overview of the gaze distribution in a scene ($M_1 = 3.79$, $SD_1 = 0.83$; $M_2 = 4.00$, $SD_2 = 1.22$) and that it is useful to define *arbitrary viewpoints* for them ($M_1 = 4.05$, $SD_1 = 0.69$; $M_2 = 3.75$, $SD_2 = 1.09$). The opinions about the usefulness of *object-based attentional maps* varied considerably among the first group, resulting in an almost uniform distribution ($M_1 = 3.32$, $SD_1 = 1.34$). The qualitative feedback suggests that the principle of this technique was not clear to everybody, which prohibited a correct assessment of its usefulness. The second group, which could try out this technique, agreed that object-based attentional maps give a quick overview about objects' visual attractiveness and how they are spatially related to each other ($M_2 = 3.88$, $SD_2 = 0.6$). *Surface-based attentional maps* are considered capable of providing a more detailed overview on how the visual attention is spread across a model's surface ($M_1 = 3.58$, $SD_1 = 0.99$; $M_2 = 4.25$, $SD_2 = 1.30$). Participants moderately agreed that different *heat signatures* are important ($M_1 = 3.53$, $SD_1 = 0.99$; $M_2 = 2.88$, $SD_2 = 0.78$). A *combination of attentional maps* for data exploration at different levels of detail has been rated as important ($M_1 = 3.47$, $SD_1 = 0.78$; $M_2 = 4.00$, $SD_2 = 0.50$).

In conclusion, the techniques were considered useful for examining gaze distributions in 3D VEs. Thereby, the quantitative results from both groups did not vary significantly. The visualization experts helped to identify insufficient descriptions for object-based attentional maps in the online survey.

3.2.2 Subjective Impressions & Suggested Improvements

In general, we received very positive feedback from both groups about the novel gaze visualizations. Especially the *surface-based attentional map* brought forward new ideas and was described as "interesting for studying the way people perceive a virtual object". Some of the most common suggestions and impressions are reported below.

One suggested improvement was the data integration from several users. It was also proposed that it would be desirable to incorporate viewing directions in the *surface-based attentional map*. The *projected attentional map* could be represented as a flow map with vector representations or as a series of semi-transparent attention clouds.

Several eye tracking experts remarked that "visualizations are for quick inspection, numbers are for statistics". We partly agree, since visualizations may reveal critical relationships in substantial data sets that can later be validated using statistical techniques. Thereby, 3D gaze data is even more difficult to evaluate than data collected for 2D stimuli. In addition, gaze visualizations are highly valuable for conveying results to customers and colleagues.

4. DISCUSSION AND CONCLUSION

In this paper we have introduced three novel aggregated gaze visualizations for application in 3D VEs: *projected*, *object-based*, and *surface-based attentional maps*. These techniques allow studying gaze data in such environments with three levels of granularity, making them a flexible toolset for rapid analysis of gaze data. Our results indicate the high potential for facilitating usability studies of 3D user interfaces for HCI researchers and practitioners.

However, visual gaze analysis of 3D VEs is still in an early stage and therefore offers much potential for further development and future work. Besides the improvement and optimization of the presented techniques, it is highly desirable to represent data from several users. In addition, possibilities to combine various data, such as fixation data and viewing directions, have to be investigated.

5. ACKNOWLEDGMENTS

This work was partially supported by the European Community FP6 FUGA research project (NEST-PATH-028765), as well as the BMBF (Bundesministerium für Bildung und Forschung) funded ViERforES project. We would like to thank all participants in our user study.

6. REFERENCES

[1] Duchowski, A.T. A breadth-first survey of eye-tracking applications. Behavior Research Methods, Instruments, and Computers, Vol. 34, No. 4, (2002), 455-470.

[2] Duchowski, A.T., Medlin, E., Cournia, N., Gramopadhye, A., Melloy, B. and Nair, S. 3D Eye Movement Analysis for VR Visual Inspection Training. ETRA '02, ACM Press (2002), 103-110.

[3] Lessing, S. and Linge, L. IICap: A new environment for eye tracking data analysis. Master's thesis. University of Lund, Sweden, 2002.

[4] Möller, T. and Trumbore, B. Fast, minimum storage ray/triangle intersection. SIGGRAPH '05, ACM Press (2005).

[5] Ramloll, R. Trepagnier, C., Sebrechts, M. and Beedasy, J. Gaze data visualization tools: opportunities and challenges. IV '04 (2004), 173–180.

[6] Stellmach, S., Nacke, L., Dachselt, R., Lindley, C. Trends and Techniques in Visual Gaze Analysis. COGAIN'09 (2009), 89-93.

[7] Velichkovsky, B.M. and Hansen, J.P. New technological windows into mind: there is more in eyes and brains for human-computer interaction. CHI '96, ACM Press (1996), 496-503.

[8] Wooding, D.S. Fixation maps: Quantifying eye-movement traces. ETRA '02, ACM Press (2002), 31-36.