Gaze-Dependent Depth-of-Field Effect Rendering in Virtual Environments

Radosław Mantiuk, Bartosz Bazyluk, and Anna Tomaszewska

West Pomeranian University of Technology in Szczecin,
Żołnierska 49, 71-210, Szczecin, Poland
rmantiuk@wi.zut.edu.pl

Abstract. This paper presents gaze-dependent depth-of-field (DOF) rendering setup, consisting of high frequency eye tracker connected to a graphics workstation. A scene is rendered and visualised with the DOF simulation controlled by data captured with the eye tracker. To render a scene in real-time, the reverse-mapped z-buffer DOF simulation technique with the blurring method based on Poisson disk is used. We conduct perceptual experiments to test human impressions caused by simulation of the DOF phenomenon and to assess benefits of using eye tracker to control the DOF effect rendering in virtual environments. Additionally, we survey the eye tracking technologies suitable for virtual environments and preview techniques of the real time DOF rendering.

1 Introduction

Depth-of-field (DOF) of the eyes is the range of distances where objects appear to be in focus for the human visual system [5]. A finite aperture of the eye maps a 3D point from the scene to a circular region in an image, called the *circle of confusion* (CoC). Overlapped CoCs make defocused objects blurred. The DOF effect improves photorealism of rendered images and supports perception of depth [6]. However, important task of this phenomenon is to make a focus object attract the human's attention. In real world, humans do not see blurred areas and to notice defocused objects some cognitive process must be arranged.

In the paper we present a gaze-dependent depth-of-field system, consisting of a high frequency eye tracker connected to graphics workstation. The hardware setup and visualisation software enable capturing observer's gaze direction, computing CoC, and rendering a 3D scene with the DOF effect. The reverse-mapped z-buffer technique with the blurring method based on Poisson disk is used to render the DOF effect in real time.

To our best knowledge, it is the first real time eye tracking setup which does not require head mounted devices or head stabilisation (e.g. chin-rest or bite-bar) to integrate eye tracking with the real time DOF visualisation. So, we are able to test humans' assessment of the DOF effect not disturbed by any devices that could divert ones attention.

We argue that gaze-dependent rendering of the DOF effect is very useful for realistic visualisation of computer graphics. In particular, it is especially

M. Ma, M. Fradinho Oliveira, J. Madeiras Pereira (Eds.): SGDA 2011, LNCS 6944, pp. 1–12, 2011.

needed in *serious games* where immersing in virtual environment (VE) plays a primary role.

In this work we conduct perceptual experiments that gather observers' judgements on presence of the DOF simulation in virtual environments. We evaluate humans' assessments concerning using the eye tracker to control the DOF rendering. We also search for the most plausible blurriness level suitable for VEs but adjusting the parameters of our optical model.

In this paper we extend the work presented by Hillaire et al. in [7] and test the gaze-dependent DOF visualisation for a more general virtual reality application, not only for first-person-player game. In contrary to [7] we evaluate the Plausibility Illusion [24], which refers to the illusion that the scenario being depicted is actually occurring, rather than the Place Illusion - a sensation of being in a real place (e.g. immersion in VE).

Additionally, we present the saliency maps achieved during experiment to assess fixation distribution in VE applications. They differ significantly from centre-oriented distributions reported in [27].

The paper starts with a survey of eye tracking techniques (Sect. 2) guided by their suitability for VEs. In Sect. 3 we discuss previous work regarding real time visualisation of the DOF effect. The implementation of our gaze-dependent depth-of-field simulation is presented in Sect. 4. Sect. 5 describes the subjective experiment, depicts its results and compares with results achieved in [7]. We conclude the paper in Sect. 6.

2 Eye Tracking Technologies for Virtual Reality Applications

Eye tracking is a technique of gathering real-time data concerning gaze direction of human eyes. In particular, position of the point (called *point-of-regard*) that a person is looking at is captured [17]. This information is acquired in numerous ways encompassing intruisive and remote techniques.

The most suitable for VE systems are remote techniques that use cameras to capture image of the eye. Even if they require some intrusive head mounted devices [8, Sect. 6], they are still acceptable for many VE applications.

The most common remote eye trackers apply the *corneal reflection* (CR) method. The eyes are exposed to direct invisible infra-red (IR) light, what results in appearance of Purkinje image with reflection in the cornea. The reflection is accompanied by image of the pupil. Captured by a video camera sensitive to the infra-red spectrum, the relative movement of both the pupil and corneal reflections are measured, what enables to estimate an observer gaze point. Commercial eye trackers can achieve the accuracy below 0.5 degree [9,10]. The CR eye trackers require calibration to estimate position of a head relatively to the screen plane. Then, it is possible to calculate the estimated screen-space gaze point coordinates with frequency higher than eye saccades[1]. In our user stud-

[1] The saccades are defined as rapid movement of the gaze-point, characteristic for the human visual system (HVS).

ies (see sections 5) we use the corneal reflection eye tracker manufactured by SensoMotoric Instruments company (SMI RED250 [9]). Detailed reviews of eye tracking techniques are presented in [11] and [8].

The eye tracking is considered to be helpful human-computer interface, especially suited for serious games applications. For example, it is intended to be used in diagnosing of attention deficit disorder (ADD) [4, Chapter 10], or to develop games that allow researchers to capture psychophysiological data, reflecting how players perceive digital games [18]. In this paper we focus on the quality of rendering and serious games are an example of using VEs in which the realism of rendering is particularly important. We argue that the DOF effect should be taken into consideration during visualisation of a serious game environment because it may significantly affect the impact of game on the player. In [1] the depth of field is mentioned as a method of enhancing visual realism of surgical simulators. The need of the DOF rendering in cultural heritage tasks is reported in [2]. In [3] importance of the DOF rendering for flight simulators is highlighted. Serious games should follow the quality of graphics available in the latest computer games, particularly in relation to natural phenomena.

3 Depth-of-Field Rendering in Virtual Environments

In a physical camera, the rays cast on the sensor are subject to refractions produced by physical properties of the lens. They appear in photography as the phenomena named circles of confusion (CoC), with diameter varying according to the ratio of the aperture to the relative distance of the object from the camera. As they blend together, image appears blurred for objects that are closer or farther from the lens focus plane. The portion of a scene that appears acceptably sharp in the image is called the depth-of-field (DOF).

An extended camera model rather than the idealised pinhole camera is needed in computer graphics to simulate the DOF effect [12]. The common method for DOF rendering is to render multiple images with varying view location (camera location) while keeping the point of focus fixed. However, this accumulation technique comes at a high cost of multi-pass rendering. Better performance can be achieved using *reverse mapping* methods [14,13]. The mentioned DOF algorithms suffer from artefacts such as intensity leakage (colour bleeding from in-focus objects onto blurred background), depth discontinuity and occlusion problem at silhouette edges of objects. Recently, advanced techniques were proposed that generate results comparable to off-line DOF rendering based on ray tracing. For detailed review of the real time DOF rendering methods refer to [19] and [22, Sect. 10.13].

In our demo application we implemented the reverse mapping technique [13]. We introduced some improvements to this algorithm to decrease visibility of artefacts (see Sect. 4 for details). We used this simple technique to favour the high rendering speed rather than the accuracy of DOF visualisation. Our main goal was to achieve real time rendering synchronised with eye tracker output.

Previous experiments with depth of field effect controlled by an eye tracker are described in [7]. They involved the reverse-mapped z-buffer method, together with minor artefacts correction. It though suffered from a drawback typical to classic depth-buffer approaches that calculate the blurriness with circle of confusion modelling. They are subject to heavily disturbing depth discontinuity problem when an out-of-focus object occludes the in-focus background. It results in hard edges of blurred object silhouette (see Figure 1, right).

In Hillaire et al. [7] the perceptual experiments were conducted to estimate uers' immersion in VE. They report the increase of immersion feeling for the visualisation setup equipped with an eye tracker. However, they use ASL 6000 eye tracker with the accuracy of less than one degree of visual angle together with a chin-rest to maintain participant's head at the same position. In our opinion, this solution strongly influences the immersion feeling and could distort results of the experiment. Moreover, Hillaire et al. project is based on the first-person-player game engine. The cognitive processes during game-play differ from those experienced during walk through a typical virtual environment. They reduce acuteness of how an observer experiences the DOF visualisation. This problem is even stronger in [15] where eye tracker is not used and a participant is assumed to look at the centre of the screen.

In [7] a technique called auto-focusing was implemented to help translating the gaze point coordinates provided by an eye tracker into actual focus distance. We find this techniques similar to estimation of fixation points based on space and time sampling. They use rectangular region surrounding the gaze point and objects' weights to average saccade movement and sample saccades with 15 Hz frequency.

In our solution, we use a high accuracy eye tracker and an improved algorithm of the DOF rendering. We conduct perceptual experiments based on a typical VE application.

Fig. 1. Left: our solution blurs not only the object interior but also the object silhouette. Right: the depth discontinuity problem, the silhouette of the candlestick remains sharp.

4 Gaze-Dependent Depth-of-Field Rendering

DOF rendering. Our implementation of the depth of field algorithm is an extension to the reverse-mapped z-buffer technique with the blurring method based on Poisson disk samples [13,14]. The algorithm requires two rendering passes. On the first pass a scene is rendered, outputting the colour as well as depth and blurriness factor. The factor is estimated by the CoC derived from the thin-lens model:

$$CoC = a \cdot \left|\frac{f}{d_0 - f}\right| \cdot \left|1 - \frac{d_0}{d_p}\right|, \tag{1}$$

where a is a diameter of the lens aperture, f is the focal length of lens, d_0 is the distance between the focus plane and the lens (objects located in this plane are sharp), and d_p is a distance from an object point to the lens.

The d_0 is equal to the depth value of a pixel corresponding to the current gaze-point captured by the eye tracker. The depth is taken from z-buffer and transformed to world-space. Finally, the offset (60cm) is added to d_0 to take into consideration distance between the monitor screen and observer's eyes.

On the second pass, the image from the first pass is filtered with a variable-sized filter kernel to simulate the circle-of-confusion. The filter kernel consists of 15 samples and its size is adjusted on a per-pixel basis using the value of CoC for the central sample read from a pre-generated, screen-size texture. The samples are then averaged to derive the final blurred colour.

To address the depth discontinuity problem we follow the solution presented in [16] where CoC value is calculated using both originally computed version and its blurred counterpart. As a result, the blur spreads outside the silhouette of a defocused object forming soft edges. The drawback of this method is blurring of background in object's vicinity (see Figure 1, left). Time consuming techniques like multi-view rendering [20] or per-pixels layers [21] techniques should be implemented to solve this problem.

Implementation. In our OpenGL 3.2-based DOF renderer the scene is rendered to a buffer (FBO) with two textures attached: one for colour output, and the second for depth values storage. Then the camera distance from the gaze-point (captured by the eye tracker) is acquired and assumed to be the focus distance. The CoC_o values are calculated with Equation 1 based on the depth values and current focus distance. The values are duplicated and stored in two separate textures: the nearer-than and farther-than the focus plane. The „near" texture is down-sampled to 1/8th of its size in every dimension, in order to obtain the blurred CoC_b values. To obtain the final CoC which determines the actual blurring radius, the following equation is used:

$$CoC = 2 \cdot max(CoC_o, CoC_b) - CoC_o. \tag{2}$$

Finally, the image is displayed on a screen (for the complete diagram see Fig. 2).

The program was implemented in *C#* language, using *OpenGL 3.2* library and *GLSL 1.50* shader language. We utilised the *.Net* based *OpenTK 1.0* library as the application's graphics abstraction layer.

Eye tracking. To query physical eye tracker and acquire screen-space coordinates of the current user gaze point, we implemented independent library utilising its own thread. The library bases on the API delivered with SMI RED250 eye tracker, which tracks an observer gaze direction and returns its value transform to the screen coordinates. The gaze point position is passed to the renderer and is used to compute the focus plane.

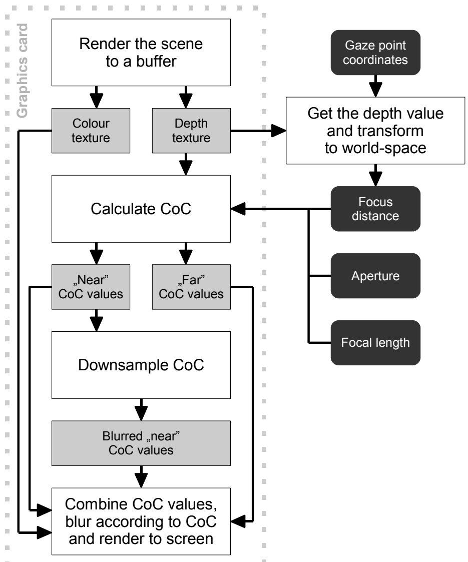


Fig. 2. The gaze-dependent depth-of-field rendering algorithm controlled by data from eye-tracker

5 Experimental Evaluation

To study humans' preferences regarding visualisation of DOF in VEs we conducted experiments based on the application which renders an example virtual environment with the DOF effect controlled by the RED250 eye tracker.

5.1 Participants

21 participants (20 males and one female) took part in the experiment, however only 16 finished it. 5 persons were not able to pass the eye tracker calibration and validation process because of severe errors in the returned data. The participants age was between 21 and 24 years, with average of 21.81 years. They had normal or corrected to normal visual acuity and correct colour vision tested with the Ishihara charts. All participants had basic expertise in imaging (they passed basic computer graphics course). None of them was aware of the technical details of the experiment. They were informed that the goal of the experiment is to judge realism of DOF simulation.

5.2 Stimuli

The rendered scene presents a fantasy-world interior of a magician's house (example screenshots are presented in Figure 3). We built a scene with complexity corresponding to environments used in simple VE systems (the scene consists of over 21 thousand triangles and 14 high resolution textures). We select two different animation sequences (Figures 3c and 3d) and two static scenes (Figures 3a and 3b) showing different parts of the scene. Each animation lasted 15 seconds and a static scenes was exposed for 8 seconds. We decided to suppress participants to navigate in the scene to pay their attention on judgement of the DOF quality rather than controlling VE. However, the animations were rendered in real time and taking control over a camera was possible.

Every scene was rendered in four different ways: without eye tracker control (focus depth and its changes were predefined for each scene) and with full eye tracker control over the DOF visualisation for three different blur levels (aperture $a = 19$, $a = 7$, and $a = 2$ (meaning of the a parameter is explained in equation 1)).

The experiments were run on a 22" Dell E2210 LCD display (1680x1050 pixel resolution, 475 mm screen width and 298 mm screen height) offering good colour and contrast reproduction. Observers sit in the front of the display in 60-70 cm distance. The illumination in the laboratory was subdued by black curtains to minimise the effect of display glare and to focus observers' attention on the VE visualisation. The scene was rendered with 60 fps on a PC with 2.8 GHz Intel i7 930 CPU and NVIDIA Geforce GTX 480 graphics card.

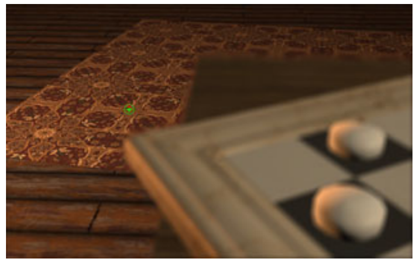
5.3 Experimental Procedure

Observers were asked to read a written instruction before every experiment. Following [23] recommendation, the experiment started with a training session in which observers could familiarise themselves with the task, interface, and the eye tracker operation. After that session, they could ask questions or start the main experiment. To ensure that observers fully attend the experiment, two random trials were shown at the beginning of the main session without recording the results. The scenes were displayed in a random order and with a different randomisation for each session. Two consecutive trials showing the same animation were avoided if possible. No session took longer than 8 minutes to avoid fatigue.

The experiment started with the eye tracker calibration. This procedure took about a minute and consisted in observation of the markers displayed in different areas of the screen. The relation between known marker positions and observer's gaze points are used to compute the gaze point in the screen coordinates (in pixels). To assure correct calibration, the validation was executed after the calibration. Each experiment session with the validation error over 60 pixels (roughly 1.5 degrees of the visual angle) was aborted .

In the next step, the actual experiment was applied. We based it on the *single stimulus* experimental technique. The scenes were displayed one by one separated

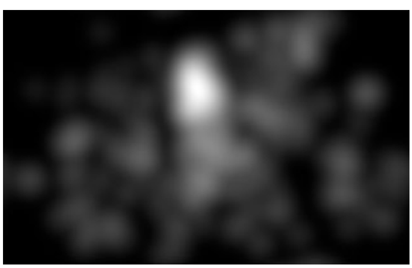
(a) Candle From Above



(b) Checkboard Corner

(c) Fireplace Through Candelabrum

(d) Floor Sweep



(e) Saliency map for Candle From Above

(f) Saliency map for Checkboard Corner

Fig. 3. Example screenshots from the virtual reality renderer. The saliency maps for a) and b) are presented in the bottom row.

by a grey screen with the slider and question: *"Does the presented depth of field effect simulation look realistic?"* . Observers were asked to use this slider and judge the quality of the DOF effect on a continuous scale from *Very realistic*, through *Realistic*, *Fairly realistic*, *Unrealistic* to *Very unrealistic*. The procedure was repeated for every scene.

At the end of each session the observer was asked to fill a questionnaire with questions concerning his age, sight condition and his opinions regarding visualisation of DOF in virtual environments.

5.4 Results

Figure 4 depicts results of the experiment. The participants disliked the DOF simulation without eye tracker control (mean $zscore = -0.78$, $sem = 0.11$). The scenes with eye tracker control and medium blur ($a = 7$) were judged as the best ones (mean $zscore = 0.54$, $sem = 0.07$). The remaining conditions achieved comparable results: mean $zscore = 0.11$ ($sem = 0.1$) for $a = 2$ and mean $zscore = 0.15$ ($sem = 0.11$) for $a = 19$.

ANOVA does not reveal significant difference of the results for various scenes ($p = 0.48$) and observes ($p = 0.99$). The results are also not significantly dependent on the accuracy of eye tracker validation.

In Figures 3e and 3f saliency maps for two example scenes are presented. These maps depict combined fixation points for every observer. For the free-walking tasks in VE we notice the observers' tendency to fixate their attention at the screen centre (reported in [27]), however distribution of saliency maps is strictly depend on a scene content. Moreover, the top-down visual attention scheme (assessment of quality of the DOF effect) induces observers to switch their gazes between near and far objects.

5.5 Discussion

In the experiment we decided to evaluate the Plausibility Illusion (Psi) rather the Place Illusion (PI) [25,24]. The latter one defines a sensation of being in a real place (immersion in VE) which was difficult to reproduce in our experimental setup, even with the remote eye tracker utilisation. On the contrary, observers were able to assess whether the virtual simulation of DOF is comparable to the natural phenomenon and judge a basic concept of the Psi - things happen in the way they should happen.

Our results show the same trends as reported in [7]: the use of eye tracker has a crucial influence on the naturalness of the DOF visualisation.

We found that the accuracy of eye tracker below an acceptable threshold (below 1.5 degrees in our case) is an important factor influencing the results. We noticed that the deviations of gaze positions in the screen plane make the computation of the focus plane difficult, especially for thin objects located in closer planes. In our opinion, the non-intruisive eye tracking with accuracy comparable to the HVS is still a challenging problem.

During the experiment we noted participants' opinions concerning the VE rendering with DOF simulation. Three of them found annoying and verbally reported the lack of eye tracker control for some scenes. The accommodation time to a new focus plane seemed to be too short for some observers. It is an interesting observation since we set this time to 370 ms according to state-of-the-art physiological studies . Participants also reported problems with focusing objects moving in the close planes. Implementation of the smooth pursuit detection should help to solve this issue.

The quality of scene rendering has significant influence on impression of the DOF naturalness. We found rendering without visible artefact more important than complexity of the shading model. A human can accept low quality graphics but artefacts divert his attention and disrupt the DOF quality.

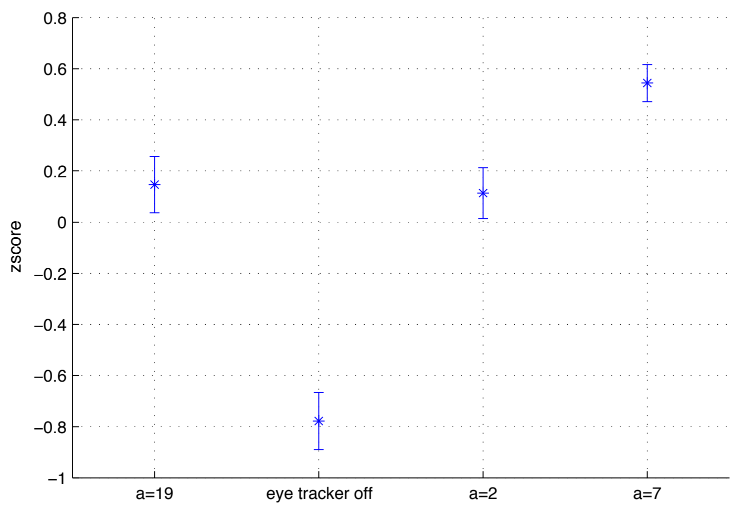


Fig. 4. Result of experiments showing observers' assessment of visualisation quality of the DOF effect simulation. The circles depict the mean zscore for four different DOF configurations: with eye tracker for three blur levels from the strongest ($a = 19$), through barely visible ($a = 2$) to medium ($a = 7$) and without eye tracker (eye tracker off). Error bars show the standard error of mean (SEM) of the zscores measured for every observer and every scene. Zscore values correspond to Likert scale values from Very realistic (positive values) to Very unrealistic, respectively.

6 Conclusions and Future Work

We have presented the depth-of-field visualisation technique in which depth of the focus plane is controlled by the eye tracker device. We conducted perceptual experiments to evaluate humans' impressions regarding to existence of the DOF phenomenon in the VE systems. The results suggest that people notice and prefer the DOF visualisation controlled by eye tracker. The best impression was achieved with the medium blurriness level.

In our hardware setup the eye tracker had limited accuracy (distortions up to 1.5 degrees) which influenced correctness of the blur location and caused blur flickering. To remove this artefact, further studies are needed to improve computations of the fixation point [26]. Development of alternative non-intruisive eye tracking techniques of higher accuracy seems to be an alternative solution.

Acknowledgements. This work was supported by the Polish Ministry of Science and Higher Education through the grants no. N N516 193537 and no. N N516 508539 .

References

1. Marks, S., Windsor, J., Wünsche, B.: Evaluation of game engines for simulated surgical training. In: Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia, GRAPHITE 2007, pp. 273–280 (2007)
2. Anderson, E.F., McLoughlin, L., Liarokapis, F., Peters, C., Petridis, P., Freitas, S.: Serious Games in Cultural Heritage VAST 2009: 10th International Symposium on Virtual Reality, Archaeology and Cultural Heritage - VAST-STAR, Short and Project Proceedings, pp. 29-48 (2009)
3. Myers, S.: Streamlining Simulation Development using a Commercial Game Engine Camber Corporation, Technical Report RTO-MP-MSG-069
4. Ritterfeld, U., Cody, M., Vorderer, P.: Serious Games: Mechanisms and Effects, 1st edn. Routledge, New York (2009)
5. Goldstein, E.B.: Sensation and Perception, 5th edn. Brooks/Cole Publishing Company (1998)
6. Mather, G.: The use of image blur as a depth cue. Perception 26, 1147–1158 (1997)
7. Hillaire, S., Lecuyer, A., Cozot, R., Casiez, G.: Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments. In: Proc. of IEEE Virtual Reality, pp. 47–50 (2008)
8. Duchowski, A.T.: Eye Tracking Methodology: Theory and Practice, 2nd edn. Springer, London (2007)
9. RED250 Technical Specification. SensoMotoric Instruments GmbH (2009)
10. Tobii T/X series Eye Trackers. Product Description. Tobii Technology AB, 2nd edn. (2009)
11. Morimoto, C.H., Mimica, M.: Eye gaze tracking techniques for interactive applications. Computer Vision and Image Understanding 98(1), 4–24 (2005)
12. Demers, J.: Depth of Field: A Survey of Techniques GPU GEMS. Addison-Wesley, Reading (2004)
13. Riguer, G., Tatarchuk, N., Isidoro, J.: Real-time depth of field simulation. haderX2: Shader Programming Tips and Tricks with DirectX 9.0, 529–579 (2002)
14. Potmesil, M., Chakravarty, I.: Modeling motion blur in computer-generated images. ACM SIGGRAPH Comput. Graph. 17(3), 389–399 (1983)
15. Hillaire, S., Lecuyer, A., Cozot, R., Casiez, G.: Depth-of-field blur effects for first-person navigation in virtual environments. In: Proc. of the ACM Symposium on Virtual Reality Software and Technology, pp. 203–206 (2007)
16. Hammon, E.: Practical post-process depth of field. GPU Gems 3, Hubert Nguyen, NVIDIA Corporation (2008)

17. Poole, A., Ball, L.J.: Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects. Encyclopedia of Human-Computer Interaction, C. Ghaoui, Idea Group, Inc., Pennsylvania (2005)
18. Sasse, D.: A Framework for Psychophysiological Data Acquisition in Digital Games Master's thesis (2008)
19. Lee, S., Eisemann, E., Seidel, H.P.: Real-Time Lens Blur Effects and Focus Control. ACM Transactions on Graphics, SIGGRAPH 2010 (2010)
20. Lee, S., Eisemann, E., Seidel, H.P.: Depth-of-Field Rendering with Multiview Synthesis. ACM Trans. Graph (SIGGRAPH ASIA 2009) 28(5), 1–6 (2009)
21. Lee, S., Kim, G.J., Choi, S.: Real-Time Depth-of-Field Rendering Using Splatting on Per-Pixel Layers. Computer Graphics Forum 27(7), 1955–1962 (2008)
22. Aknine-Moller, T., Haines, E., Hoffman, N.: Real Time Rendering, 3rd edn. A K Peters, Stanford (2008)
23. ITU-R.REC.BT.500-11: Methodology for the subjective assessment of the quality for television pictures (2002)
24. Slater, M., Spanlang, B., Corominas, D.: Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. ACM Trans. Graph. 29(4), 92:1–02:9. (2010)
25. Slater, M.: Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philosophical Transactions of the Royal Society B: Biological Sciences 364(1535), 3549–3557 (2009)
26. Hillaire, S., Lécuyer, A., Regia-Corte, T., Cozot, R., Royan, J., Breton, G.: A real-time visual attention model for predicting gaze point during first-person exploration of virtual environments. In: Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology (VRST 2010), Hong Kong, pp. 191–198 (2010)
27. Kenny, A., Delaney, H., Mcloone, S., Ward, T.: A preliminary investigation into eye gaze data in a first person shooter game. In: Proceedings of European Conference on Modelling and Simulation. Addison-Wesley, Reading (2005)