Simulated foveated depth-of-field blur for virtual reality systems

Razeen Hussain
razeen.hussain@edu.unige.it
Fabio Solari
fabio.solari@unige.it
Manuela Chessa
manuela.chessa@unige.it

Department of Informatics, Bioengineering, Robotics, and Systems Engineering,
University of Genova,
Italy

The emergence of modern VR devices facilitates new and novel experiences for users above and beyond what is possible with traditional audio-visual displays. However, its widespread usage has been hindered because users tend to feel discomfort after its prolonged usage. This discomfort or visual fatigue occurs due to the differences between visual experience in the real world and the virtual world. Incorporating spatial blur while providing stereoscopic 3D stimuli has shown to reduce visual fatigue [4]. Spatial blur can be introduced by foveated rendering [1] or by depth-of-field (DoF) effects [2].

In this paper, we develop a system that combines DoF rendering with gaze-contingent foveated rendering. All processing is done pixel-wise at the shader level in the linear colour space using image space methods in order to have real-time performance.

For both effects, a 1D Gaussian filter was used for introducing the blur. A two pass shader was used to implement the filter. In the first pass, the filter was passed in the vertical direction and the output was stored on a frame buffer. In the second pass, the filter was applied in the horizontal direction and the resulting frame was output to the HMD. A kernel of length 15 pixels was used for both passes of the Gaussian filter.

For DoF effects, a depth map was created using a depth texture. A Z-buffer stores the depth information of all the vertices in the scene. Objects in the scene that are at the accommodative distance were kept in high acuity while the smoothing filter was applied to the other regions. The amount of blur depends on how far each object is from the plane of fixation. Blur can be defined as the diameter of the circle of confusion CoC over which a distant point is imaged at the retina when the lens is focused at another distance [3]. The size of CoC can be seen as a representation for the parameter of the smoothing filter, i.e., the standard deviation of the Gaussian filter is directly related to the diameter of the circle of confusion ($\sigma_d \propto CoC$). Using this assumption, σ_d can be expressed as:

$$\sigma_d = K \left| \frac{1}{d_0} - \frac{1}{d_1} \right| \tag{1}$$

where d_0 is the depth of the pixel under fixation, d_1 is the depth of the pixel being rendered and parameter K is the fitting of the aperture A, the posterior nodal distance s and the constant relating CoC and σ_d.

For the multi-region foveation, the overall view is divided into three sections, namely the foveal, near and mid peripheral regions. The reference center of the image is the fixation point and all regions are drawn around it. The central region is output to the frame buffer without any further processing. A Gaussian filter is applied to all the remaining pixels with σ_f depending on the location of the pixel. For the near and mid peripheral regions, σ_f was defined as 1.7 and 3.5 pixels respectively.

A blending function was incorporated into the system to remove artifacts arising on the transitional regions. We define regions as $i = 1,2,3$ with 3 representing the innermost region. The blending function $B(x,y)$ is defined as:

$$B_i(x,y) = \begin{cases} 0 & R(x,y) \leq R_i \\ \frac{R(x,y)-R_i}{R_{i-1}-R_i} & R_i < R(x,y) < R_{i-1} \\ 1 & R(x,y) \geq R_{i-1} \end{cases} \tag{2}$$

where $R(x,y)$ is the distance between the rendered pixel coordinates and the pixel coordinates of the fixation point, and R_i and R_{i-1} are the radii of the transitional regions where $R_i < R_{i-1}$. The magnitudes of these radii depend on the resolution of the HMD and on visual eccentricity.

As the rendered pixel moves closer to the inner circle, the value of the blending function approaches 0 and likewise approaches 1 when the pixel is closer to the outer limit of the transitional region. For pixels where the blending function is between 0.0 and 1.0, the output image is given by:

$$O(x,y) = B_i(x,y)I_i(x,y) + (1 - B_i(x,y))I_{i-1}(x,y) \tag{3}$$

Figure 1: Example output from the foveated depth-of-field effect. Point of fixation is on the vase.

where $I_i(x,y)$ and $I_{i-1}(x,y)$ are the outputs from the Gaussian filters from i^{th} and $(i-1)^{th}$ regions. This way a percentage from each level is taken to form the output in the transitional regions.

To combine the two effects, we compute σ for both at each pixel but only use the minimum of the two values for the Gaussian filter.

For qualitative evaluation, we conducted a user study on depth perception in order to better understand the influence of using foveated DoF blur effects. The objective was to investigate whether the blur effects help perceive scene depth better and to get a preliminary evaluation of the system for visual discomfort. In a pilot study, 12 persons participated in 3 evaluation sessions each. Each session had 15 trials with blur disabled and 15 trials with blur enabled. Objects of various sizes and shapes were randomly placed on a table in a virtual scene. The reference object was indicated with a bright yellow spotlight to draw attention of the user. The users were given 4 seconds to observe the scene, then they were asked "how many objects are at the same depth of the reference one?". The subjects were then asked to indicate their answer by selecting a number on a virtual keypad integrated into the scene. After completing the experiment, the subjects were asked to fill a subjective questionnaire in order to evaluate their experience with using the system.

We observed that the performance either improved considerably or stayed the same. User performance did not deteriorate for any subject. Most of the users were overestimating the objects at the same scene depth, i.e., they gave a higher answer than the true value. We observed a 27% reduction in the error of the perceived objects by incorporating the foveated depth-of-field effect. Generally, users found the transitions smooth and did not perceive any noticeable artifacts in the foveal region. None of the users reported any discomfort while using the system.

We believe that our system can be useful in reducing visual fatigue in virtual reality systems and can be resourceful in mitigating the vergence-accommodation conflict in VR. As a next step, a thorough simulator sickness study with an already integrated eye-tracking module will be carried out to evaluate the system's effectiveness in reducing visual fatigue.

[1] B. Bastani, E. Turner, C. Vieri, H. Jiang, B. Funt, and N. Balram. Foveated pipeline for ar/vr head-mounted displays. *Information Display*, 33(6):14–35, 2017.

[2] K. Carnegie and T. Rhee. Reducing visual discomfort with hmds using dynamic depth of field. *IEEE Computer Graphics and Applications*, 35(5):34–41, 2015.

[3] R. T. Held, E. A. Cooper, J. F. O'Brien, and M. S. Banks. Using blur to affect perceived distance and size. *ACM Trans. Graph.*, 29(2): 19:1–19:16, 2010.

[4] G. Maiello, M. Chessa, F. Solari, and P. J. Bex. Simulated disparity and peripheral blur interact during binocular fusion. *Journal of Vision*, 14(8):13–13, 2014.