GeoGCD: Improved Visual Search via Gaze-Contingent Display

Kenan Bektaş
Zurich University of Applied Sciences
Centre for Aviation
kenan.bektas@zhaw.ch

Arzu Çöltekin*
University of Applied Sciences and Arts Northwestern Switzerland
Institute for Interactive Technologies
arzu.coltekin@fhnw.ch

Jens Krüger
University of Duisburg-Essen and University of Utah
High Performance Computing
jens.krueger@uni-due.de

Andrew T. Duchowski
Clemson University
Visual Computing
duchowski@clemson.edu

Sara Irina Fabrikant
University of Zurich
Department of Geography
Digital Society Initiative
sara.fabrikant@geo.uzh.ch

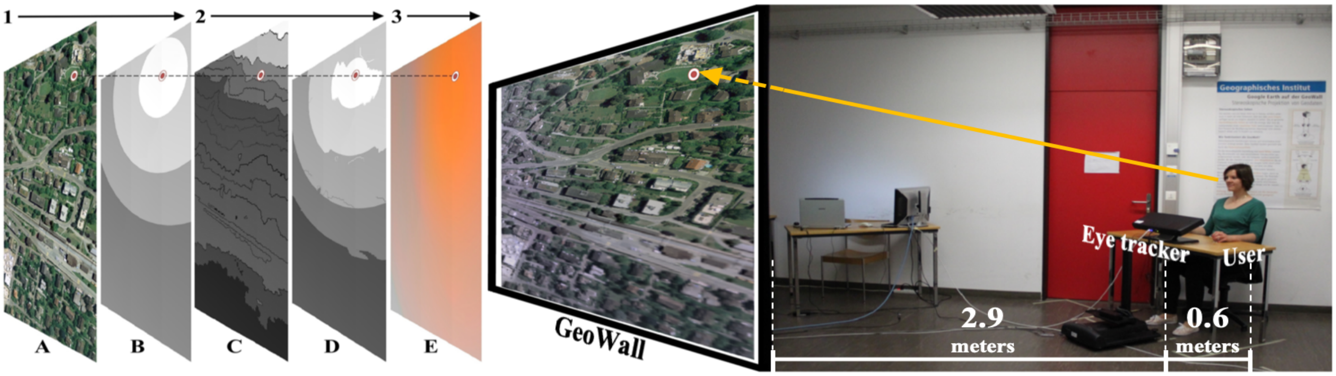

Figure 1: A combination of three visual perception models degrades resolution and color from the point of interest (top right) toward out-of-focus regions. First, the input image (A© swisstopo) is processed with a contrast sensitivity function (CSF in B). Then, the depth-of-field (DOF) simulation introduces an additional blur based on the digital elevation model (DEM in C) (as illustrated in D). Finally, a color degradation mask (E) is applied. Left part of this figure reproduced from Bektaş et al. [2015] by kind permission of the © Eurographics Association 2015. The image on the right © Kenan Bektaş.

ABSTRACT

Gaze-Contingent Displays (GCDs) can improve visual search performance on large displays. GCDs, a Level Of Detail (LOD) management technique, discards redundant peripheral detail using various human visual perception models. Models of depth and contrast perception (e.g., depth-of-field and foveation) have often been studied to address the trade-off between the computational and perceptual benefits of GCDs. However, color perception models and combinations of multiple models have not received as much attention. In this paper, we present GeoGCD which uses individual contrast, color, and depth-perception models, and their combination to render scenes without perceptible latency. As proof-of-concept, we present a three-stage user evaluation built upon geographic image interpretation tasks. GeoGCD does not impair users' visual search performance or affect their display preferences. On the contrary, in some cases, it can significantly improve users' performance.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; • **Computing methodologies** → **Rasterization**; **Perception**;

KEYWORDS

gaze-contingent displays, contrast, color, depth perception, depth-of-field simulation, visual crowding, visual search

ACM Reference Format:
Kenan Bektaş, Arzu Çöltekin, Jens Krüger, Andrew T. Duchowski, and Sara Irina Fabrikant. 2019. GeoGCD: Improved Visual Search via Gaze-Contingent Display. In *Eye Tracking for Spatial Research (ET4S@ETRA'19), June 25–28, 2019, Denver , CO, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3317959.3321488

1 INTRODUCTION AND BACKGROUND

In the context of gaze-based interaction, the foveated, or gaze-contingent display (GCD) [Duchowski 2018; Duchowski and Çöltekin 2007] retains high-resolution information only within the foveal parts of an image and discards detail in the periphery. Similar to view-dependent simplification [Luebke and Erikson 1997], or Level Of Detail (LOD) management [Koulieris et al. 2014; Luebke et al. 2002], GCDs can degrade spatial, temporal, or chromatic detail [Duchowski et al. 2009; Duchowski and Eaddy 2009; Geisler and Perry 1998b, 2002]. Pixel-based GCDs use MIP-mapping [Duchowski 2004], Laplace filtering [Böhme et al. 2006], or coarse

---

ET4S@ETRA'19, June 25–28, 2019, Denver , CO, USA

ACM ISBN 978-1-4503-6730-1/19/06…$15.00
https://doi.org/10.1145/3317959.3321488

pixel shading [Patney et al. 2016; Vaidyanathan et al. 2014] to decrease peripheral resolution. Model-based GCDs, on the other hand, can have substantial computational benefits in rendering virtual scenes by reducing geometric complexity in the periphery [Guenter et al. 2012; Levoy and Whitaker 1990; Luebke et al. 2000]. While managing LOD, GCDs mimic human vision and respond to gaze without necessarily affecting the selection of objects.

The Human Visual System (HVS) concurrently processes contrast, color, and depth information, with highest acuity in the fovea [Resnikoff 1989; Valois 2000]. In vision research, such properties of the HVS have been modeled by **Visual Perception Models** (VPMs), among which, chromatic VPMs seem to have been exploited least [Duchowski and Çöltekin 2007]. Evaluation of a combination of multiple VPMs appears to be missing from the literature. Compared to individual models, a combined model could generate a foveated image that would be more perceptually similar to a retinal image. A combined model might also lead to an optimized LOD management while concurrently discarding spatial and chromatic details from the out-of-focus parts of a GCD [Bektaş et al. 2015].

Besides the VPMs, it is important to consider the context in which a GCD would be used. One particular task, *visual search* is relevant for many visual displays. Studies on visual search offer insights into how humans make decisions and solve problems [Eckstein 2011; Geisler et al. 2006]. Researchers have tested GCDs with abstract or synthetically generated stimuli to study the effect of gaze-contingent viewing on human visual search performance [Loschky and McConkie 2000; Parkhurst et al. 2000, 2001; Shioiri and Ikeda 1989; Watson et al. 1997]. In some cases, GCDs can improve visual search performance, possibly by masking peripheral visual distractors and reducing visual complexity [Murphy and Duchowski 2007; Murphy et al. 2009]. In other cases, GCDs might attract, guide, or enhance the capacity of users' attention [Loschky and McConkie 2002; Nikolov et al. 2003; Toet 2006]. A GCD should be indistinguishable from a uniform resolution display, provided that the underlying VPMs are implemented in a perceptually plausible way [Reingold et al. 2003].

Objects in the periphery are hard to recognize in the presence of nearby objects (i.e., flankers or distractors), due to *visual crowding* [Levi 2008; Louie et al. 2007; Pelli et al. 2004]. According to Strasburger et al. [2011], "crowding is one of the key characteristics that distinguish peripheral from foveal vision". In foveal vision, the effect of crowding is limited, whereas in peripheral vision it occurs at large distances [Levi 2008]. Bouma's [1970] Law explains the spatial extent of crowding or the interaction between a flanker and a target as a function of the eccentricity of the target [Strasburger et al. 2011]. According to Bouma's formulation, "For a stimulus at X° eccentricity, an open distance of roughly 0.5X° is required for complete isolation" [Bouma 1970]. In crowding-related studies, letters or Gabor patches [Chung et al. 2001] or human face silhouettes (e.g., Louie et al. [2007]; Murphy and Duchowski [2007]; Murphy et al. [2009]) are used as targets and flankers. However, except for a few examples [Murphy and Duchowski 2007; Murphy et al. 2009; Schneider et al. 2011a,b], researchers have not examined the effect of gaze-contingent viewing on visual crowding, especially with natural scenes such as aerial images.

An important activity afforded by geographic imagery is *image interpretation*, which inherently includes visual search. For instance, in time-critical search and rescue or surveillance situations, trained personnel search for specific targets (e.g., a person or a vehicle) over a series of images. According to Eckstein [2011], the "knowledge of target is mentioned as key to successful search" by image interpretation experts (i.e., radiologists who scrutinize X-rays or analysts who search intelligence-related targets in satellite images). At times, even for experts, the image interpretation is time-consuming—like finding a needle in the haystack—because the scanned content is often information-rich. Unless the viewed image is familiar, the full scene, or possibly large parts of it, must be scanned to find the elements of interest [Netzel et al. 2017]. In such cases, a GCD could offer some help. However, empirical studies assessing user experience with GCDs on visually complex natural scenes such as in geographic image interpretation are rare.

Most previous user evaluations with GCDs have been conducted on desktop displays (e.g., Duchowski et al. [2009]; Parkhurst et al. [2001]) or head mounted displays (HMDs) [Padmanaban et al. 2017; Stengel and Magnor 2016; Watson et al. 1997]. However, GCDs might offer additional benefits with larger displays [Geisler and Perry 1998a; Ware 2013]. Large displays provide a human-scale environment and their design requires a human-centric perspective [Andrews et al. 2011]. Empirical evidence suggests that increasing display size can improve users' orientation and visual search performance [Ball and North 2005; Ball et al. 2005; Tan et al. 2003]. Large displays such as GeoWalls are used in research and education [Batty 2008]. However, only a few studies have tested space-variant visualizations with large displays [Baudisch et al. 2002], and the gaze-contingency paradigm is rarely studied on large displays.

1.1 Contributions

We offer three original contributions. First, we present the first user evaluation of GCDs that systematically compares participants' visual search performance with multiple VPMs. This yields insights into how VPMs need to be adjusted on a large display, such as a GeoWall. As a result, we present *the Adjusted COMBO* model as the second contribution. Third, we provide a detailed discussion about the implications of gaze-contingent search with our VPMs from several points of view, including the stimuli, display size, and task.

2 GAZE-CONTINGENT IMAGE PROCESSING

Bektaş et al. [2011; 2012; 2015] have developed prototype geographic GCDs. We present a standalone geographic GCD (GeoGCD) based on OpenGL and shader programming (GLSL) implementation of three VPMs and their combinations. A live demonstration of the GeoGCD was presented by Bektaş and Çöltekin [2018].

In fragment shaders, for each pixel of an input image, a new intensity can be MIP-mapped with respect to a weighted Euclidean distance between that pixel and the point of interest, or POI. In our first model, **CSF** weighting relies on a contrast sensitivity function defined by Geisler and Perry [1998a]. This model reflects the sensitivity of the human eye to contrast changes in natural viewing conditions Bektaş et al. [2015]; Murphy and Duchowski [2007]; Wang and Bovik [2001]. To create a gaze-contingent visualization, the intensity of a pixel at a certain eccentricity is sampled from the LOD that has the lowest possible resolution level that is not perceptually distinguishable by the human eye at that eccentricity.

```
// Interpolated coordinates of a bounding box for input images
in vec2 UV;
// GLFW samplers for the aerial image, DEM and color mask
uniform sampler2D myAerialImage, myDEM, myRGBMask;
// Current point of interest detected by the eye tracker
uniform float poi_x , poi_y;
// Luminance vector for high-definition displays
vec3 L = vec3 (0.212 , 0.715 , 0.072);
// Output pixel
out vec3 pixel;

void main() {
   // 2D foveation based on Geisler and Perry [1998a]
   csf_Level = getCSFMipLevel (poi_x , poi_y);
   // DOF simulation based on Rokita [1996]
   dof_Level = getDoFLevel (poi_x , poi_y, myDEM);
   // Mipmapping for spatial degradation
   I = texture2D ( myAerialImage, UV, dof_Level + csf_Level );

   // Align and scale color degradation mask
   maskUV = alignMask ( poi_x , poi_y );
   // Sampled color mask
   M = texture2D ( myRGBMask , maskUV );

   if( (0.0, 0.0) <= maskUV.xy <= (1.0, 1.0) )
   {
      // Color degradation based on Duchowski et al. [2009]
      // for all pixels overlapping with the scaled color mask
      pixel = vec3 (I * M) + dot(L, (I *( 1.0 - M )));
   }
   else {
      pixel = vec3 (0.0 , 0.0 , 0.0) + dot(L, I);
   }
}
```

Listing 1: Procedure for spatial (CSF and DOF) and chromatic degradation (COLOR) with COMBO.

Based on the suggestions of Geisler and Perry [1998a], a maximum of four LODs are computed.

Next, the **DOF** model is a depth-of-field simulation, wherein the intensities are sampled according to the normalized spatial distances between the near and far planes of a given scene. A thin-lens model proposed by Rokita [1996] is easily applicable for real-time gaze-contingent processing. Empirical evidence suggests that such models increase perceived realism [Hillaire et al. 2008; Mantiuk et al. 2011] even if the scenes were not rendered in real-time [Mauderer et al. 2014]. Therefore, Rokita's model was used in the implementation of the GeoGCD. The GeoGCD calculates spatial depth information in real-time based on a Digital Elevation Model (DEM), which is analogous to a z-buffer used in 3D graphics. However, beyond approximately 23 meters viewing distance, such as in the aerial images, the HVS cannot distinguish near and far objects [Smith and Atchison 1997]. Thus our DOF model linearly augments the amount of blur introduced in aerial images.

Recently, Mauderer et al. [2016] used GCDs for the manipulation of perceived color to enhance color discrimination. One particular study, by Duchowski et al. [2009], used *color zone maps* [Sakurai et al. 2003] to generate a trichromatic visual mask. This mask can simulate peripheral chromatic sensitivity of the human eye to red, green, and blue light. The algorithm used by Duchowski et al. is easy to implement, and their findings provide a good baseline for the user evaluation of the GeoGCD. In our third model, **COLOR**, chromatic LOD is maintained through a color degradation mask that is adapted from Duchowski et al. [2009]. Provided that the color degradation mask and input image are correctly aligned for a particular point of interest, the eccentricity-based chromatic degradation is obtained in two steps: First, the intensities in the RGB channels of every pixel in the input image (I) are scaled with the values in the corresponding pixel of the mask (M) by a point-wise multiplication operation. The result of scaling is an image with degraded color and luminance towards the periphery. Second, in order to recover the luminance information, the RGB values of each pixel in the output image are interpolated between the original pixel intensities and a luminance vector (L) that is used in high-definition displays.

The **COMBO** model, a combination of CSF, DOF, and COLOR, takes a uniform resolution image, a corresponding rasterized DEM, the color degradation mask, and the POI as input (see Listing 1). COMBO performs gaze-contingent image processing in three consecutive stages as illustrated in Figure 1. In the first stage, COMBO calculates an index `miplevel` based on the CSF and reduces image resolution as a function of eccentricity. Then, using DOF and the DEM, COMBO computes another index `dof` for all pixels based on their distance to the focal plane. At this point, the pixels that remain within the foveal region and depth-of-field preserve their resolution. Other pixels—those that remain within the foveal region but outside the depth-of-field, and those in the peripheral and out-of-focus regions—are assigned lower resolutions. See Figure 1 and note the additional blur introduced by the DOF. At the final stage, the COLOR model adjusts the chromaticity of each pixel, as described by Duchowski et al. [2009].

The GeoGCD can handle various image formats (e.g., `bmp, jpg, tiff`) and multiple (geo-referenced) vector overlays (i.e., shapefiles). With minor modifications, the GeoGCD and VPMs are applicable for a much broader class of problems (e.g., medical data visualization). Thus, the use of the GeoGCD is not limited to geographic purposes. In the next section, we present a benchmark on the GeoGCD's system performance with two hardware setups.

2.1 GeoGCD System Performance

For seamless gaze-contingent viewing, Loschky and Wolverton [2007] suggest a 60 ms system latency—a threshold below which

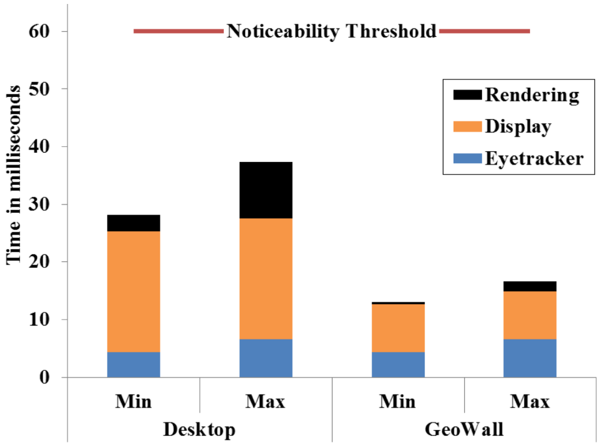

Figure 2: In Desktop and GeoWall setups, the minimum and maximum computation time required by the GoeGCD (for rendering, displaying, and eye tracking) remain below the 60ms noticeability threshold reported in Loschky and Wolverton [2007] and Albert et al. [2017].

Table 1: Average participant performance and subjective ratings per visualization type. The best score in visual comfort and confidence is 5, and in noticeability the best score is 0. The last row shows the results of one-way repeated measures ANOVA.

Display	TCT	Accuracy	Visual Comfort	Confidence	Noticeability
Uniform	6.9± 0.2 sec.	88%± 2	4.2± 0.1	4.4± 0.05	2.1± 0.2
CSF	7.2± 0.1 sec.	90%± 3	4.1± 0.1	4.5± 0.1	2.4± 0.1
COLOR	7.2± 0.1 sec.	90%± 1.5	3.9± 0.1	4.3± 0.03	2.4± 0.1
DOF	7.4± 0.2 sec.	88%± 3	4.2± 0.1	4.4± 0.05	2.0± 0.1
COMBO	7.3± 0.1 sec.	93%± 2	4.1± 0.01	4.3± 0.05	2.3± 0.1
ANOVA	F = 1.644 $p = .18$, $\eta^2 = .13$	F = .611 $p = .657$, $\eta^2 = .05$	F = 2.259 $p = .078$, $\eta^2 = .17$	F = 1.284 $p = .291$, $\eta^2 = .105$	F = 1.533 $p = .209$, $\eta^2 = .122$

latency can be assumed to be unnoticeable, depending on task. Albert et al. [2017] corroborated this finding reporting latency requirements for gaze-contingent rendering in virtual reality (VR).

The overall latency of a GCD depends on the eye-tracking latency (*EL*) and the time required for rendering. The total rendering time can be further decomposed into the time required by the GPU (*RT*) and the monitor or display latency (*DL*). We considered the GeoGCD's latency in two setups, Desktop and GeoWall, and with images at *low* (1280 × 800) and *high* (1920 × 1080) resolution.

We used a Tobii TX300 eye tracker that samples a pair of gaze coordinates every 3.3 milliseconds (ms) at 300 Hz and that needs an additional 1.0 to 3.3 ms to record the data [Tobii 2010]. Thus, total *EL* remains between 4.3 to 6.6 ms.

The Desktop setup includes a personal computer: Intel Core i5 2.67 GHz Quadcore CPU, 8 Gb RAM, NVIDIA GeForce GT 520 GPU, and a 60 Hz Phillips LCD monitor. The monitor requires 21 ms to scan out a new frame from the GPU.

The GeoWall setup includes a workstation: Intel Core i73960x 3.30 GHz CPU, 32 GB RAM, NVIDIA GeForce Quadro 4000 GPU, and a large back projection display with DepthQ HDs3D-1 Stereo Projector, which operates at 120 Hz and 8.3 ms total latency.

The total latency of the Desktop setup ($EL + DL + RT$) is between 4.3 + 21 + 2.8 = 28.1 ms and 6.6 + 21 + 9.7 = 37.3 ms for the low- and high-resolution images, respectively. Similarly, the total latency of the GeoWall setup varies between 4.3 + 8.3 + 0.4 = 13 ms and 6.6 + 8.3 + 1.7 = 16.6 ms for the low- and high-resolution images, respectively. These results, as illustrated in Figure 2, show that the GeoGCD can render a gaze-contingent frame in less than 60 ms.

3 USER EVALUATION OF GEOGCD

We designed three studies to gauge the effects of the VPMs and their combination on visual search performance and on user preference when using the GeoWall. The first two pilot studies informed the third, main study. In the first pilot study, spatial parameters of each VPM were set to default values as reported in earlier work [Duchowski et al. 2009; Geisler and Perry 1998a]. In the second pilot study, parameters of the VPMs were adjusted to reduce visible artifacts observed in the first pilot study. The outcome of the second pilot study yielded the Adjusted COMBO model, evaluated in the main study. All studies followed a within-subjects design. Ethics approval was provided by the University of Zürich.

3.1 Pilot Study 1

The independent variable of this pilot study was *visualization type* at five levels: uniform resolution display (**Uniform**) as the control, and four gaze-contingent displays rendered with **CSF**, **COLOR**, **DOF**, and **COMBO** models. We hypothesized that with all gaze-contingent displays, visual search performance would match or exceed that of the Uniform model (**Hypothesis 1**). We measured performance along two metrics: *task completion time (TCT)* and the *accuracy* of responses. We also hypothesized that visual comfort, confidence, and noticeability ratings would be similar between the uniform and gaze-contingent displays (**Hypothesis 2**). We measured these subjective ratings along a five-point Likert scale ranging from *strongly disagree* to *strongly agree.*

Stimuli and Apparatus. We prepared two aerial images at 1:5000 scale with urban or rural content. The urban image contained many small-sized elements (e.g., roads, buildings), thus, it contained greater visual complexity than the rural image composed of large planar areas (e.g., forest and agricultural fields). Target objects were simple, identical red circles, randomly superimposed as vector overlays on the aerial images, subtending 0.55° visual angle.

We conducted the study with the GeoWall setup. The stimuli were back projected onto a 243.5×136 cm screen at a 3.5 meter viewing distance (Figure 1). Gaze was tracked with the Tobii TX300.

Participants. Twenty-one participants took part in the study. Data collected from nine participants was incomplete because they leaned forward or to the side. Only the data from 12 participants (four females, aged 25 to 38) were included in the analysis. Half of the participants had corrected vision, and none had color vision problems. Participants were compensated with a lunch coupon.

Procedure. After obtaining written consent, we informed all participants about the procedure, which consisted of two consecutive sessions lasting 70 minutes in total. Each session started with training followed by a nine-point eye-tracking calibration.

In the first session, participants went through 100 trials. A block of 10 search tasks was followed by a block of 10 count tasks. In the search blocks, the task was to find the target and click on it with the mouse. In the count blocks, participants counted the total number of targets, clicked on any location on the display, and reported the result verbally. To avoid learning effects, the visualization type, the image complexity, and the target location (for search) or the number of targets (for count) were fully counter-balanced in a Latin square design. In each trial, the TCT—the time between the start of a trial and the mouse click—was recorded in milliseconds by the GeoGCD.

In the search tasks, we measured the accuracy of responses based on the location of the mouse click. In the count tasks, we compared the reported total number of targets with the correct total.

To test our second hypothesis in the second session, we opted to use the count task instead of the search task because of its longer duration, allowing participants to notice potential visual artifacts. Using only one task type reduced the duration of this session and allowed focused engagement by the participants. Through 20 trials, all factors (i.e., visualization types, images, and number of targets) were fully counter-balanced in a Latin square design. After each trial, participants rated their level of agreement with statements related to visual comfort, confidence, and noticeability. At the end of the study, we debriefed and compensated the participants.

Results. The results of participants' task performance and subjective ratings are summarized in Table 1.

For each visualization, the average TCT is approximately 7.2 seconds, and participants spent on average 1.6 seconds per search task and 5.6 seconds per count task. Visualization had no observable effect on TCT or accuracy. For each visualization, an average subjective rating was calculated as the mean of all trials. Visualization had no statistically significant effect on perceived visual comfort, response confidence, or noticeability of visual artifacts.

Brief Discussion. Results indicate that discarding perceptually redundant spatial details (CSF, DOF, COMBO) or peripheral color degradation (COLOR, COMBO) had no observable effect on task performance (TCT, accuracy), supporting our first hypothesis.

In Pilot Study 1, the total number of trials was high. In order to limit the experiment duration, we used simple red circles as target objects which could be easily searched and counted. Thus, the participants completed the search tasks with 100% accuracy in a very short time. The overall performance results were very close for all tested visualization types. We think that our experimental design (i.e., the tasks and the properties of the target objects) was not sufficiently challenging to show potential effects of the VPMs on the performance of participants. This outcome can be interpreted as a *ceiling effect.* Thus, in the main study, we decided to select complex target objects which require more spatial-thinking.

Similarly, gaze-contingent viewing had no effect on visual comfort or confidence, and visual artifacts were not any more noticeable than in the Uniform display, also supporting our second hypothesis. Although noticeability ratings did not differ significantly, participants noticed some visual artifacts in all visualization types, including the Uniform display. In debriefing, some participants referred to reflections of blinking LEDs from the apparatus, which might explain why the noticeable artifacts existed even in the Uniform display (last column in Table 1). In the next study, potential light sources were covered with opaque or antireflective material. Because overall noticeability ratings were higher than expected, we suspected some issues with the initial adjustment of the VPMs or system latency. In the second pilot study, we explored what other factors might have caused noticeable artifacts with our VPMs.

3.2 Pilot Study 2

Creating a perceptually plausible GCD requires adjustment of the VPMs according to the experimental conditions (e.g., stimuli, display, etc.) [Parkhurst and Niebur 2002; Peli 1990; Reingold et al. 2003]. The VPMs we used in the first study were not previously user tested with complex natural scenes (e.g., aerial images). They were also not adjusted for gaze-contingent viewing on large displays at a long distance (e.g., GeoWall). A controlled readjustment of these VPMs for the GeoWall could potentially contribute to overcoming noticeability-related problems encountered in the first pilot study.

In the second pilot study, we systematically adjusted the spatial extent of the CSF and the COLOR models to reduce the noticeability of visual artifacts reported earlier. These adjustments are methodologically similar to the studies presented by Murphy and Duchowski [2001] and Guenter et al. [2012]. In the first pilot study, the average noticeability with the DOF was less than other visualization types. For this reason, the DOF was excluded.

Participants. Four graduate students (three female, aged 28 to 33) participated in this study. None reported any vision-related problems, and they were not involved in the first study.

Procedure. After obtaining written consent, we first trained participants with *mouse-contingent* visualizations displayed with the CSF and COLOR models. Because space-variant rendering was decoupled from participant's gaze, participants were expected to notice peripheral resolution and color degradation. Then, we wanted to know whether the noticeable artifacts in gaze-contingent viewing were the result of system latency or due to an inadequate adjustment of the underlying VPM. For this reason, we defined two tasks in which gaze-contingency was enabled. First, in the *restricted viewing* task, participants were asked to focus on the crosshairs that were superimposed at the center of an aerial image. Second, in the *free viewing* task, participants were asked to freely view the same image. For the free-viewing trials, we expected both the system latency and the VPMs to play a role in the noticeability. In the restricted viewing trials, system latency would *not* be a factor, because the gaze-contingent, high-resolution inset was stationary.

After a nine-point eye-tracking calibration, each participant performed 24 trials. In 12 trials, the maximum LOD of the CSF model varied as 4 (default), 5, and 6. In the remaining 12 trials, the size of the color degradation mask in the COLOR model varied from small (default) to medium and large. The parameters were varied in an ascending order and with two repetitions, in four consecutive blocks. After each trial, participants were asked whether they noticed any visual effect similar to those demonstrated during the mouse-contingent training.

Results. In the *mouse-contingent* training, all participants noticed the peripheral effects caused by the CSF and COLOR models, as expected. They identified the effect caused by the CSF as "slightly changing sharpness", or "change in blur", and the effect of the COLOR as "reduced or fading color", or "less color". This finding confirms that they could identify the two expected visual effects and would be able to respond to the questions in the study.

In restricted viewing, peripheral color degradation with the default small-sized mask (20° of the visual field in full color) was noticeable in seven of eight trials. With medium and large masks, peripheral color degradation was noticeable only in one of eight trials. The medium-sized mask degrades more color in the periphery than the large mask, without causing more noticeable artifacts. On the other hand, the medium-sized mask preserves more color in the periphery than the small mask and causes fewer noticeable artifacts. The responses to the CSF model showed that, at level 4,

Table 2: Average participant performance and subjective ratings per display. The last row in each table shows results of a one-way repeated measures ANOVA and Friedman test, respectively.

Display	Search TCT	Count TCT	Search Accuracy	Count Accuracy
Uniform	4.1± .5 sec.	9.2± .5 sec.	94%±11.1	60.3%±19.4
COMBO	4.1± .5 sec.	9.5± .5 sec.	95%±10.2	56%±17.8
Adj.COMBO	3.7± .5 sec.	**8.6± .5** sec.	98.3%±5.1	54.3%±22.2
ANOVA	F = 1.948 $p = .15$, $\eta^2 = .05$	F = 5.589 $\boldsymbol{p = .005}$, $\eta^2 = .13$	F = 2.492 $p = .09$, $\eta^2 = .062$	F = 1.447 $p = .242$, $\eta^2 = .037$

Display	Image Quality	Visual Comfort	Confidence	Noticeability
Uniform	3.90± .6	4.4± .6	4.2± .74	1.9± .8
COMBO	4.02± .7	4.4± .6	4.1± .77	1.9± .6
Adj.COMBO	3.97± .7	4.2± .6	4.0± .73	2.0± .7
Friedman-ANOVA	χ^2 (2) = 0.182, $p = .913$	χ^2 (2) = .667, $p = .717$	χ^2 (2) = 1.351, $p = .509$	χ^2 (2) = 1.341, $p = .511$

the peripheral resolution reduction was noticeable in five of eight trials. At level 5, the same effect was noticeable in three of eight trials. At level 6, the results were not different from those at level 5.

In the free-viewing task, participants noticed color degradation in all trials with the small mask. Color degradation was noticeable in three of eight trials with the medium mask, and only once with the large mask. In five of the 24 trials, participants noticed some flicker effect on the display. Following the restricted viewing results, it is perceptually more plausible to use the medium mask in the COLOR model. With the CSF model, only one participant noticed blur in the periphery in all trials. The other three participants reported no noticeable blur at any level.

The maximum LOD of the CSF is proportional to its computational efficiency. Compared to level 4, level 5 introduces more blur to an image, which requires less memory after lossy compression. From a perceptual standpoint, in restricted viewing, level 5 introduced fewer noticeable artifacts than level 4. Thus, it appears that level 5 is more plausible for use than level 4, from both computational and perceptual perspectives.

Given these observations, the adjusted versions of the CSF (maximum LOD = 5) and the COLOR (28.4° mask size) models were combined with the DOF model to yield the *Adjusted COMBO* model.

3.3 Main Study

The main study compared the Adjusted COMBO model against both the COMBO and the Uniform models. As in the first pilot study, we hypothesized that with the GCD, participants' performance would match or exceed that of the Uniform display (**Hypothesis 1**). We further hypothesized that participants' ratings of image quality, visual comfort, confidence, and noticeability would be similar between Uniform and the GCD (**Hypothesis 2**). In the main study, we measured and analyzed dependent variables (i.e., performance measures and subjective ratings) as in the first pilot study. We evaluated user performance based on TCT and accuracy. We measured the subjective ratings along a five-point Likert scale from *strongly disagree* to *strongly agree.*

Stimuli and Apparatus. We tested three displays: Uniform, COMBO, and the Adjusted COMBO. For each of the search and count tasks, six aerial images and corresponding DEMs were shown at a 1:1000 scale. To avoid potential learning effects, horizontally and vertically flipped versions of the images were used for the two gaze-contingent displays. For each search task, an aerial image was shown along with a target snippet—a small portion extracted from each image that includes an object with salient spatial features (e.g., a landmark or crossroads). The target snippet subtended 5° visual angle. Each target snippet was presented in the center of a gray background covering the entire screen (Figure 3). The aerial image was also shown in full-screen mode (Figure 1). For the count tasks,

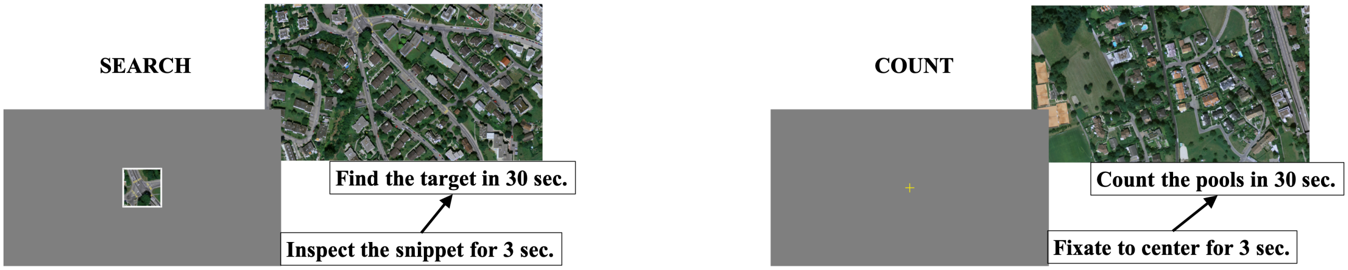

Figure 3: The first two search and count task blocks of the first session of the main study (Aerial images © swisstopo).

swimming pools with similar color and shape were used as target objects. The total number of targets per image changed from 4 to 10. In all images, the targets were positioned arbitrarily and were of variable size (0.8° and 1.2°). The main study was conducted with the same experimental apparatus as in the first pilot study (GeoWall, workstation, and eye tracker).

Participants. Thirty-nine participants (11 female, 16 with corrected vision, aged 23 to 45) were recruited as domain experts from the Remote Sensing and GIScience units in the Department of Geography of the University of Zürich. All participants were offered refreshments and a lunch coupon for their participation.

Procedure. After obtaining participants' written consent, we briefed each on the procedure and trained each with two search and two count tasks. The main study was conducted in two consecutive sessions (60 minutes in total). After a nine-point eye-tracking calibration, participants went through 36 trials: a block of six search tasks was followed by a block of six count tasks.

In the search trials, the target snippet was first shown for three seconds. Participants then had a maximum of 30 seconds to find the target within the aerial image and click on it with the mouse. In the count trials, participants started by focusing on the crosshairs shown in the center of the screen for three seconds. They then had 30 seconds to count the total number of swimming pools in the image, click the mouse, and report their answer verbally. Figure 3 illustrates the procedure of the first session.

After a short break, the second session started again with a nine-point eye-tracking calibration. Participants were assigned only three count tasks. In each trial, the GeoGCD presented one of the three display types in random order. After each trial, participants filled in a questionnaire that included statements related to image quality, visual comfort, confidence, and noticeability. At the end of the main study, we debriefed and compensated the participants.

Results. Participants' performance is summarized in Table 2.

Descriptive statistics show that participants searched fastest with the Adjusted COMBO display, whereas their performance with the Uniform and COMBO displays was similar. Participants also found the location of the target most accurately with the Adjusted COMBO display, followed by the COMBO and the Uniform displays. However, these differences were not statistically significant, and thus the display had no effect on overall search task performance (i.e., speed or accuracy).

Participants counted fastest with the Adjusted COMBO display, followed by the Uniform and COMBO displays. A Kolmogorov-Smirnov test showed that only the data from the COMBO was normally distributed ($p = .092$), and thus a log transformation was applied to all data. Mauchly's test indicated that the assumption of sphericity had not been violated ($p > .05$), $\chi^2(2) = 2.845, p = .241$. A one-way repeated-measures ANOVA with a Greenhouse-Geisser sphericity correction ($\varepsilon = .931$) showed that the average TCT for the count tasks was affected by display, $F(2, 76.0) = 5.589$, $p = .005$, $\eta^2 = .13$. Pairwise comparisons with Tukey's post hoc test showed a significant difference between the Adjusted COMBO and Uniform displays ($p = .026$) and between the Adjusted COMBO and COMBO displays ($p = .006$). No significant difference was found between the Uniform and COMBO displays ($p = .219$). The overall accuracy of counting tasks was not affected by display type. In summary, the Adjusted COMBO significantly improved participants' counting performance compared to the Uniform and COMBO displays.

Results of subjective ratings are summarized in Table 2. Participants' ratings of the overall image quality, visual comfort, confidence, and noticeability of the visual artifacts were not affected by display type. In both the first pilot study and in the main study, the overall noticeability of the visual artifacts was the same across all tested display types. In comparison to the first pilot study, the average noticeability ratings with the Uniform and COMBO models were reduced by 5% and 9%, respectively. This reduction indicates an improvement in the perceptual plausibility of the GeoGCD, as participants noticed fewer visual artifacts.

4 DISCUSSION

The results of the main study support our hypothesis that the gaze-contingent displays with the COMBO and Adjusted COMBO models *do not impair* performance in image interpretation tasks, despite reducing most of the peripheral spatial and chromatic detail in real-time. In fact, the Adjusted COMBO *improved* the time required for counting the swimming pools. Furthermore, participants equally preferred the tested displays (i.e., gaze-contingent and uniform). We believe they expressed this preference mainly because they did not appear to notice the peripheral detail reduction, which we interpret as evidence of successful GCD implementation in terms of both system latency and VPMs.

An important finding is improved task performance with the Adjusted COMBO. Compared to the Uniform display, the Adjusted COMBO afforded 0.6 seconds faster counting performance, on average. Our analysis showed that the speed-up in the time required for counting is statistically significant. This improvement can be attributed to several factors, namely the Visual Perception Models, the size of the display and respective viewing distance, and the given tasks. Below we summarize our interpretations of how each of these factors may have played a role.

Visual Perception Models. The size of the high resolution inset in a GCD can affect users' visual search performance in three ways. Reducing the inset size below 2° *deteriorates* visual search performance [Parkhurst and Niebur 2002]. An inset of 4.1° to 5° can, however, simulate normal viewing conditions on desktop displays. With an HMD, a 30° gaze-contingent inset leads to similar user performance with a uniform resolution display [Watson et al. 1997]; thus, the effect is *neutral.* Third, Murphy et al. [2009] reported a mid-sized (10°) gaze-contingent inset that *improved* visual search performance on a desktop setup, attributing this improvement to the mid-sized inset's ability to mask peripheral distractors. We also believe that peripheral masking with a GCD is the main reason behind performance improvement observed with our Adjusted COMBO model. Because the Adjusted COMBO model discards more spatial detail in the periphery than the COMBO model while preserving more chromatic detail than the COMBO, in effect, the Adjusted COMBO model reduces the visual crowding effect, largely due to spatial (not chromatic) detail. Performance improvement was observed only in the count task, which takes longer than a search task as counting inherently involves multiple searches. We believe that during counting, the improvement with each search accumulates over time. Our observations with the Adjusted COMBO (i.e., the improvement in TCT without impairing accuracy) provide evidence on the perceptual soundness of gaze-contingent viewing. However, the results of our user evaluation and the findings reported in previous work must be compared and interpreted with caution because of the differences in the experimental conditions (e.g., VPMs, parameter adjustment and display type). In this work, we included only three fundamental VPMs and their combinations with a specific parameter adjustment. We believe, it is necessary to gather further empirical evidence (i.e., from controlled studies on parameter adjustment of the color and depth perception models, and combinations of multiple models) to determine the suitability of GCDs in use-cases outside the controlled laboratory environments.

Display Size. In desktop setups, eye trackers are often either integrated or closely positioned at the physical display [Tobii 2010]. In our GeoWall setup, the eye tracker was positioned about 3 meters away from the display. This long viewing distance affected the eye tracking calibration, and thus the GeoGCD's functionality. Without reliable gaze input, a GCD cannot be correctly aligned to its user's Point Of Interest (POI), and peripheral detail reduction becomes apparent to the user, which is undesirable. In addition to the calibration issues, our setup was vulnerable to some distractors, which are less likely in a desktop setup. Our test environment was isolated from potential external audio and visual distractors, but due to the long viewing distance (Figure 1), blinking LEDs and light reflections from the floor were noticed by participants in the first pilot study. The adjustments of the VPMs and the optimization of the experimental setup eliminated most of the visual distractions prior to the main study. The Adjusted COMBO can be thought of as a GeoWall compatible version of the COMBO model, which originally relied on individual models customized to a desktop setup. Thus, performance improvement with the Adjusted COMBO may be linked to the size and viewing distance of the physical display. To simulate normal viewing conditions, or to improve user performance with a GCD, as observed with our GeoGCD, the interplay between display size and the spatial extent and type of VPM must be carefully considered. Our results obtained from our GeoWall setup suggest that GCDs may be effective at large viewing distances and on large displays, provided that distractions unrelated to the task are minimized.

Task Dependence. The observed improvement in the visual search performance is potentially meaningful in image interpretation tasks that require scrutinizing targets among a series of images e.g., aerial photographs. A consecutive examination of 100 images for counting certain targets would take approximately 920 and 860 seconds with the Uniform and Adjusted COMBO, respectively. Thus, with the Adjusted COMBO, image interpretation can be finished 60 seconds earlier, without compromising accuracy. Within this extra time, seven additional images can be inspected.

4.1 Limitations of the User Evaluation

Our research examines some of the bottom-up mechanisms of visual perception. The cognitive, or top-down, effects of gaze-contingent viewing, such as its relation to visual working memory, are left for future work. Experts in radiology or photogrammetry build an intuitive understanding of targets and distractors over time. Thus, they can solve visual search tasks in a reflexive way and perform better than a novice can. In our main study, participants were asked to memorize an object that was presented in a preview snippet. However, we did not study how gaze-contingent visual search performance was affected with respect to the participants' expertise in image interpretation, or their memory capacity.

Optimal visual search performance is strongly linked to eye movement behavior Geisler et al. [2006]; Parkhurst and Niebur [2002]. To better understand how visual search strategies are affected by gaze-contingent viewing, eye movement events such as fixations, saccades, and pupillary activity need to be studied. The eye movement data that was collected during our evaluation of the GeoGCD needs to be prepared for further analysis and made public. Such analyses can potentially provide researchers additional and vital insights regarding visual search strategies and any performance benefits observed with GCDs, such as the Adjusted COMBO model.

5 CONCLUSION

Our novel GCD implementation with multiple VPMs, the GeoGCD, serves as a tool to study the potential of the gaze-contingent paradigm on large displays. The findings presented in this paper shed some light on the user experience with gaze-contingent display of aerial images. However, the GoeGCD and the VPMs can easily be adopted to many other visualization contexts as well. We provide our VPMs as fragment shaders and encourage their reuse with other GCDs for further evaluation.

5.1 Future Outlook

As eye tracking becomes more reliable and affordable [Feit et al. 2017], and small and light-weight eye trackers compatibly work with near-eye displays and HMDs [Orlosky et al. 2017; Padmanaban et al. 2017; Roth et al. 2016], the relevance of GCDs will continue for the visualization of natural or virtual scenes. In the future, the assessment of alternative VPMs (e.g., models of motion perception, stereoscopic viewing, peripheral and foveal masking, as well as transparency-based VPMs and semantic or content aware VPMs), which are not included in this work, would be a fruitful area of research. Investigations on the interplay between the perceptual (e.g., visual crowding [Whitney and Levi 2012]) and computational (e.g., visual masking [Ferwerda et al. 1997; Williams et al. 2003]) aspects of GCDs with various content, tasks, and user groups would provide valuable additions to the literature. For example, unmanned aerial vehicles (UAVs) have become available in the consumer market that can deliver images or videos from complex environments (e.g., in the wilderness, natural disaster) [Whitehead and Hugenholtz 2014]. Thus, it seems plausible to develop a head-mounted GeoGCD that can assist UAV pilots and payload operators in real-time image interpretation. In the long term, following rigorous usability testing with both novice and expert users, we believe the GeoGCD can be deployed in emergency response, e.g., search and rescue situations.

ACKNOWLEDGMENTS

This research is partially supported by the Swiss National Science Foundation: 200021-120434/1, and partially funded by the first author. We gratefully thank the participants of our user studies and anonymous reviewers of this paper for their valuable feedback.

REFERENCES

Albert, R., Patney, A., Luebke, D., and Kim, J. (2017). Latency requirements for foveated rendering in virtual reality. *ACM Transactions on Applied Perception (TAP)*, 14(4):25.

Andrews, C., Endert, A., Yost, B., and North, C. (2011). Information Visualization on Large, High-resolution Displays - Issues, Challenges, and Opportunities. *Information Visualization*, 10(4):341–355.

Ball, R. and North, C. (2005). Effects of tiled high-resolution display on basic visualization and navigation tasks. *CHI '05 extended abstracts on Human factors in computing systems - CHI '05*, pages 1196 – 1199.

Ball, R., Varghese, M., Carstensen, B., Cox, E. D., Fierer, C., Peterson, M., and North, C. (2005). Evaluating the Benefits of Tiled Displays for Navigating Maps. *In Proc. IASTED-HCI '05, ACTA*, pages 66–71.

Batty, M. (2008). Virtual reality in geographic information systems. In Wilson, J. P. and Fotheringham, A. S., editors, *The Handbook of Geographic Information Science*, chapter 17, pages 317–335. Blackwell.

Baudisch, P., Good, N., Bellotti, V., and Schraedley, P. (2002). Keeping Things in Context: A Comparative Evaluation of Focus Plus Context Screens, Overviews, and Zooming. *CHI '02 Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, (4):259–266.

Bektaş, K. and Çöltekin, A. (2011). An Approach to Modeling Spatial Perception for Geovisualization. *Procedia - Social and Behavioral Sciences*, 21:53–62.

Bektaş, K. and Çöltekin, A. (2012). Area of Interest Based Interaction and Geovisualization with WebGL. *The Graphical Web Conference*, pages 1–14.

Bektaş, K., Çöltekin, A., Krüger, J., and Duchowski, A. T. (2015). A Testbed Combining Visual Perception Models for Geographic Gaze Contingent Displays. In Bertini, E., Kennedy, J., and Puppo, E., editors, *Eurographics Conference on Visualization (EuroVis) – Short Papers*, pages 67–71.

Bektaş, K. and Çöltekin, A. (2018). Geogcd: Geographic gaze contingent display. In *Eye Tracking for Spatial Research, Proceedings of the 3rd International Workshop*. ETH Zurich.

Böhme, M., Dorr, M., Martinetz, T., and Barth, E. (2006). Gaze-Contingent Temporal Filtering of Video. In *ETRA '06: Proceedings of the 2006 Symposium on Eye Tracking Research & Applications*, pages 109–116, San Diego, CA. ACM.

Bouma, H. (1970). Interaction Effects in Parafoveal Letter Recognition. *Nature*, 226(5241):177–178.

Chung, S. T. L., Levi, D. M., and Legge, G. E. (2001). Spatial-frequency and contrast properties of crowding. *Vision Research*, 41(14):1833–1850.

Duchowski, A. T. (2004). Hardware-Accelerated Real-Time Simulation of Arbitrary Visual Fields. In *ETRA '04: Proceedings of the 2004 Symposium on Eye Tracking Research & Applications*, San Antonio, TX. ACM. (Poster).

Duchowski, A. T. (2018). Gaze-based interaction: A 30 year retrospective. *Computers & Graphics*, pages –. Special Section on Serious Games and Virtual Environments.

Duchowski, A. T., Bate, D., Stringfellow, P., Thakur, K., Melloy, B. J., and Gramopadhye, A. K. (2009). On spatiochromatic visual sensitivity and peripheral color LOD management. *ACM Transactions on Applied Perception*, 6(2):1–18.

Duchowski, A. T. and Çöltekin, A. (2007). Foveated gaze-contingent displays for peripheral LOD management, 3D visualization, and stereo imaging. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 3(4):1–18.

Duchowski, A. T. and Eaddy, T. D. (2009). A Gaze-Contingent Display Compensating for Scotomata. In *EuroGraphics (Short Presentations)*, Munich, Germany. EuroGraphics.

Eckstein, M. P. (2011). Visual search: a retrospective. *Journal of vision*, 11(5).

Feit, A. M., Williams, S., Toledo, A., Paradiso, A., Kane, S., and Morris, M. R. (2017). Toward Everyday Gaze Input: Accuracy and Precision of Eye Tracking and Implications for Design. *Proc. of ACM CHI 2017*.

Ferwerda, K. a., Shirley, P., Pattanaik, S. N., and Greenberg, D. P. (1997). A model of visual masking for computer graphics. *Siggraph 97*, pages 143–152.

Geisler, W. S. and Perry, J. S. (1998a). A real-time foveated multiresolution system for low-bandwidth video communication. In Rogowitz, B. E. and Pappas, T. N., editors, *Proceedings of SPIE*, volume 3299, pages 294–305. Spie.

Geisler, W. S. and Perry, J. S. (1998b). Real-time Foveated Multiresolution System for Low-bandwidth Video Communication. In *Human Vision and Electronic Imaging*, Bellingham, WA. SPIE.

Geisler, W. S. and Perry, J. S. (2002). Real-time Simulation of Arbitrary Visual Fields. In *Eye Tracking Research & Applications (ETRA) Symposium*, pages 83–153, New Orleans, LA. ACM.

Geisler, W. S., Perry, J. S., and Najemnik, J. (2006). Visual search: The role of peripheral information measured using gaze-contingent displays. *Journal of Vision*, 6(9):1–1.

Guenter, B., Finch, M., Drucker, S., Tan, D., and Snyder, J. (2012). Foveated 3D graphics. *ACM Transactions on Graphics*, 31(6):1.

Hillaire, S., Lécuyer, A., Cozot, R., and Casiez, G. (2008). Depth-of-field blur effects for first-person navigation in virtual environments. In *IEEE computer graphics and applications*, pages 47–55, New York, New York, USA. ACM Press.

Koulieris, G. A., Drettakis, G., Cunningham, D., and Mania, K. (2014). C-LOD: Context-aware Material Level-of-Detail applied to Mobile Graphics. *Computer Graphics Forum*, 3(4):41–49.

Levi, D. M. (2008). Crowding-An essential bottleneck for object recognition: A mini-review. *Vision Research*, 48(5):635–654.

Levoy, M. and Whitaker, R. (1990). Gaze-Directed Volume Rendering. In *Computer Graphics (SIGGRAPH '90)*, pages 217–223, New York, NY. ACM.

Loschky, L. C. and McConkie, G. W. (2000). User performance with gaze contingent multiresolutional displays. In *Proceedings of the symposium on Eye tracking research & applications - ETRA '00*, pages 97–103, New York, New York, USA. ACM Press.

Loschky, L. C. and McConkie, G. W. (2002). Investigating spatial vision and dynamic attentional selection using a gaze-contingent multiresolutional display. *Journal of Experimental Psychology: Applied*, 8(2):99–117.

Loschky, L. C. and Wolverton, G. S. (2007). How late can you update gaze-contingent multiresolutional displays without detection? *ACM Transactions on Multimedia Computing, Communications, and Applications*, 3(4):1–10.

Louie, E. G., Bressler, D. W., and Whitney, D. (2007). Holistic crowding: selective interference between configural representations of faces in crowded scenes. *Journal of Vision*, 7(2):24.1–11.

Luebke, D. and Erikson, C. (1997). View-Dependent Simplification Of Arbitrary Polygonal Environments. In *Computer Graphics (SIGGRAPH '97)*, New York, NY. ACM.

Luebke, D., Hallen, B., Newfield, D., and Watson, B. (2000). Perceptually Driven Simplification Using Gaze-Directed Rendering. Technical Report CS-2000-04, University of Virginia.

Luebke, D., Reddy, M., Cohen, J., Varshney, A., Watson, B., and Huebner, R. (2002). *Level of Detail for 3D Graphics*. Morgan-Kaufmann Publishers, San Francisco, CA.

Mantiuk, R., Bazyluk, B., and Tomaszewska, A. (2011). Gaze-dependent depth-of-field effect rendering in virtual environments. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6944 LNCS:1–12.

Mauderer, M., Conte, S., Nacenta, M. a., and Vishwanath, D. (2014). Depth perception with gaze-contingent depth of field. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, (Chi):217–226.

Mauderer, M., Flatla, D. R., and Nacenta, M. A. (2016). Gaze-contingent manipulation of color perception. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 5191–5202. ACM.

Murphy, H. and Duchowski, A. T. (2001). Gaze-contingent level of detail rendering. *Eurographics*.

Murphy, H. and Duchowski, A. T. (2007). Hybrid image-/model-based gaze-contingent rendering. In *Proceedings of the 4th symposium on Applied perception in graphics and visualization - APGV '07*, volume 5, page 107, New York, New York, USA. ACM Press.

Murphy, H. a., Duchowski, A. T., and Tyrrell, R. a. (2009). Hybrid image/model-based gaze-contingent rendering. *ACM Transactions on Applied Perception*, 5(4):1–21.

Netzel, R., Hlawatsch, M., Burch, M., Balakrishnan, S., Schmauder, H., and Weiskopf, D. (2017). An Evaluation of Visual Search Support in Maps. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):421–430.

Nikolov, S. G., Bull, D. R., and Gilchrist, I. D. (2003). Gaze-Contingent Multi-modality Displays of Multi-layered Geographical Maps. pages 325–332.

Orlosky, J., Kiyokawa, K., and Takemura, H. (2017). Virtual and Augmented Reality on the 5G Highway. *Journal of Information Processing*, 25(0):133–141.

Padmanaban, N., Konrad, R., Stramer, T., Cooper, E. A., and Wetzstein, G. (2017). Optimizing virtual reality for all users through gaze-contingent and adaptive focus displays. *Proceedings of the National Academy of Sciences*, 114(9):2183–2188.

Parkhurst, D., Culurciello, E., and Niebur, E. (2000). Evaluating variable resolution displays with visual search. In *Proceedings of the symposium on Eye tracking research & applications - ETRA '00*, pages 105–109, New York, New York, USA. ACM Press.

Parkhurst, D., Law, I., and Niebur, E. (2001). Evaluating Gaze-Contingent Level of Detail Rendering of Virtual Environments using Visual Search.

Parkhurst, D. J. and Niebur, E. (2002). Variable-resolution displays: a theoretical, practical, and behavioral evaluation. *Human factors*, 44(4):611–29.

Patney, A., Salvi, M., Kim, J., Kaplanyan, A., Wyman, C., Benty, N., Luebke, D., and Lefohn, A. (2016). Towards Foveated Rendering for Gaze-tracked Virtual Reality. *ACM Trans. Graph.*, 35(6):179:1–179:12.

Peli, E. (1990). Contrast in complex images. *Journal of the Optical Society of America. A, Optics and image science*, 7(10):2032–2040.

Pelli, D. G., Palomares, M., and Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of vision*, 4(12):12.

Reingold, E. M., Loschky, L. C., McConkie, G. W., and Stampe, D. M. (2003). Gaze-Contingent Multiresolutional Displays: An Integrative Review. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 45(2):307–328.

Resnikoff, H. L. (1989). *The Illusion of Reality*, volume 10. Springer US, New York, NY.

Rokita, P. (1996). Generating depth of-field effects in virtual reality applications. *IEEE Computer Graphics and Applications*, 16(2):18–21.

Roth, T., Weier, M., Hinkenjann, A., and Li, Y. (2016). An analysis of eye-tracking data in foveated ray tracing. *IEEE Second Workshop on Eye Tracking and Visualization (ETVIZ)*, pages 69–73.

Sakurai, M., Ayama, M., and Kumagai, T. (2003). Color appearance in the entire visual field: color zone map based on the unique hue component. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 20(11):1997–2009.

Schneider, N., Dorr, M., Pomarjanschi, L., and Barth, E. (2011a). A gaze-contingent reading tutor program for children with developmental dyslexia. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization - APGV '11*, number January 2011, page 117, New York, New York, USA. ACM Press.

Schneider, N., Dorr, M., Pomarjanschi, L., and Barth, E. (2011b). Improving reading capability of children with developmental dyslexia with a gaze-contingent display. *Institute for Neuro-und Bioinformatics, University of Lübeck (http://www. perceptionweb. com/abstract. cgi*, page 124.

Shioiri, S. and Ikeda, M. (1989). Useful resolution for picture perception as a function of eccentricity. *Perception*, 18(3):347–361.

Smith, G. and Atchison, D. A. (1997). The Eye. In *The Eye and Visual Optical Instruments*, chapter 13. Cambridge University Press.

Stengel, M. and Magnor, M. (2016). Gaze-Contingent Computational Displays: Boosting perceptual fidelity. *IEEE Signal Processing Magazine*, 33(5):139–148.

Strasburger, H., Rentschler, I., and Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of vision*, 11(5):13.

Tan, D. S., Gergle, D., Scupelli, P., and Pausch, R. (2003). With similar visual angles, larger displays improve spatial performance. *Proceedings of the conference on Human factors in computing systems - CHI '03*, (5):217.

Tobii (2010). Tobii TX300 Eye Tracker Product Description. Technical report.

Toet, A. (2006). Gaze directed displays as an enabling technology for attention aware systems. *Computers in Human Behavior*, 22(4):615–647.

Vaidyanathan, K., Salvi, M., Toth, R., Foley, T., Akenine-Möller, T., Nilsson, J., Munkberg, J., Hasselgren, J., Sugihara, M., Clarberg, P., et al. (2014). Coarse pixel shading. In *Proceedings of High Performance Graphics*, pages 9–18. Eurographics Association.

Valois, K. K. D. (2000). *Seeing*. Academic Press.

Wang, Z. and Bovik, a. C. (2001). Embedded foveation image coding. *IEEE transactions on image processing*, 10(10):1397–1410.

Ware, C. (2013). *Information Visualization Perception For Design*. Morgan Kaufmann, 3 edition.

Watson, B., Walker, N., Hodges, L. F., and Worden, A. (1997). Managing level of detail through peripheral degradation: effects on search performance with a head-mounted display. *ACM Transactions on Computer-Human Interaction*, 4(4):323–346.

Whitehead, K. and Hugenholtz, C. H. (2014). Remote sensing of the environment with small unmanned aircraft systems (UASs), part 1: a review of progress and challenges 1. *Journal of Unmanned Vehicle Systems*, 02(03):69–85.

Whitney, D. and Levi, D. M. (2012). Visual Crowding: a dundamental limit on conscious perception and object recognition. *Trends Cogn Sci*, 15(4):160–168.

Williams, N., Luebke, D., Cohen, J. D., Kelley, M., and Schubert, B. (2003). Perceptually guided simplification of lit, textured meshes. *Proceedings of the 2003 symposium on Interactive 3D graphics - SI3D '03*, page 113.