Eurographics Conference on Visualization (EuroVis) (2015)
J. Kennedy and E. Puppo (Editors)

A Testbed Combining Visual Perception Models for Geographic Gaze Contingent Displays

K. Bektaş[1], A. Çöltekin[1]*, J. Krüger[2], A. T. Duchowski[3]

[1]University of Zurich, Zurich, Switzerland,*Corresponding author: arzu@geo.uzh.ch
[2]University of University of Duisburg Essen, Germany, [3]Clemson University, Clemson, SC, USA

Abstract

We present a testbed featuring gaze-contingent displays (GCDs), in which we combined multiple models of the human visual system (HVS) to manage the visual level of detail. GCDs respond to the viewer's gaze in real-time, rendering a space-variant visualization. Our testbed is optimized for testing mathematical models of the human visual perception utilized in GCDs. Specifically, we combined models of contrast sensitivity, color perception and depth of field; and customized our implementation for geographic imagery. In this customization process, similarly to the geographic information systems (GIS), we georeference the input images, add vector layers on demand, and enable stereo viewing. After the implementation, we studied the computational and perceptual benefits of the studied perceptual models in terms of data reduction and user experience in geographic information science (GIScience) domain. Our computational validation experiments and the user study results indicate the HVS-based data reduction solutions are competitive, and encourage further research. We believe the research outcome and the testbed will be relevant in domains where visual interpretation of imagery is a part of professional life; such as in search and rescue, damage assessment in hazards, geographic image interpretation or urban planning.

Categories and Subject Descriptors (according to ACM CCS): H 1.2 [User/Machine Systems]: Human Factors-Human Information Processing, H 5.2 [User Interfaces]: Input devices and strategies, I.4 [Image Processing and Computer Vision]-Image displays

1. Introduction

Aerial and satellite imagery are used for a wide variety of tasks such as change detection [GOT*11], land use classification [BB05], urban planning [BRB06], emergency and rescue operations [MWS13], hazard damage assessments [SSGM04] or visual perception research [Bc14, cDBV14]. In such tasks, an expert viewer is typically expected to visually interpret a series of large images. Especially when executed in succession, such tasks are demanding on computational resources as well as humans [cR11]. Many methods were developed to handle large images as well as large image collections. For example, storing the images in multiresolution databases and transmitting the images in tiles have become standard practices [MM96, BGS00]. As technology evolved, heavy graphics processing operations were transferred to graphics processing units (GPUs); parallel processing solutions were developed [SBA14, YZP08], and most recently, cloud-based solutions have been proposed [FCD*14, GSZ14]. Despite these developments, availability of new sensors and devices – such as the unmanned aerial vehicles (UAVs), or new small dedicated imaging satellites – make the data size issues even more pronounced today [MB13, SBSS14]. Therefore, image manipulation and LOD management efforts remain relevant [YW09]. However, despite their promise in LOD management, to the best of our knowledge, no studies attempt to combine visual perception models to produce a combined visualization solution [BÇS12]. Some HVS-inspired visualization ideas have been proposed and utilized for a relatively long time in computer graphics. For example, *level of detail management* approaches show the "sensible" amount of detail based on the limitations of human vision, such as the object's distance from the viewer's position [Red01]. [LWC*02] provides an extensive review about perceptually motivated LOD management.

In this paper, we propose a testbed for gaze-contingent vi-

sualizations. Our main goal is to study the potential of *combining* the HVS-inspired visualization paradigms, especially in geographic information science domain.

2. Models and Implementation

Our testbed combines various (existing) visual perception models (VPMs) and integrates functions to combine them in pairs or multiples, and adjust relevant spatial parameters, thus, will help determining computational as well as human performance with included models and model combinations. In all documented cases of HVS-inspired visualization solutions that we reviewed, considerable data and/or bandwidth reduction was observed (e.g., [Red01,Lin04,Ç09,DBS*09]). Besides the computational benefits, presenting parts of the scene in lower resolution and in fewer colors where the HVS does not perceive much detail has been suggested to be a *perceptually lossless* data reduction technique by various researchers [Bc11,HM97].

2.1. Foveation

The first VPM we included in our testbed is known as *foveation* which was shown to be a useful technique for reducing image resolution from (or introduce blur to) the periphery, where most of the visual detail is not resolvable by the human eye [GP98, Dc07]. Gaze-contingent displays are closely related to the foveation approach, as a GCD reacts to the viewer's gaze (captured via an eye tracker) in real-time and removes the perceptually irrelevant detail in the periphery [DCM04]. Many of the foveated GCDs use a contrast sensitivity function (CSF) to adjust the level of detail change in 2D visual filed. In standard CSF models, CSF = 1 yields the maximum spatial frequency per degree that is perceivable for the eye (*mfE*). Using simple trigonometric rules, we can define the maximum spatial frequency that can be displayed on a digital screen (*mfS*) based on the viewing distance and display dotpitch. The corresponding LOD for each pixel are calculated as:

$$LOD = \frac{mfS}{mfE} \tag{1}$$

In practical applications, the LOD in equation (1) would be normalized between 0 (corresponds to the original image resolution) and an upper limit determined by the display size, input image resolution and other constraints of the system in use. Geisler and Perry (1998) suggest a value between 4 and 6 for the maximum LOD [GP98].

2.2. Color Degradation

Similar to the contrast perception, the spatio-chromatic sensitivity of human eye declines towards the peripheral visual field. The second VPM we employed uses a color mask developed by Duchowski et al. (2009), which, in accordance with the HVS, was used to degrade the chromaticity of digital images [DBS*09]. The red, green and blue channels of the color degradation mask is shown in Figure 1. We have modified this model with an update of the coefficients of the luminance vector with: vec4(0.212, 0.715, 0.072, 1.0) as suggested for modern CRT and HDTV displays.

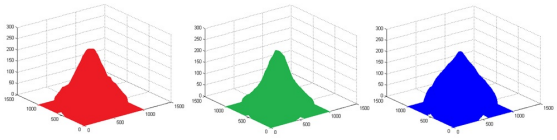

Figure 1: *Red, green and blue channels of the color degradation mask. Vertical axis represent the intensity (0-255) for each channel over (1000 X 1000) mask resolution. Note that the sensitivity to each color component is slightly different.*

2.3. Depth of Field Simulation

In three-dimensional space, simulating *depth of field* (DOF) as well as exploiting the spatial distribution of the resolution based on *stereoscopic acuity* have been suggested several times [Ç09,Lin04,DHG*14]. However, the proposed approaches have rarely been tested. We selected the thin-lens system that is proposed in [Rok96] for the DOF simulation and implemented it as the third VPM in our testbed. We integrated the model so that the testbed is responsive to the changes in the viewing distance and the pupil diameter in real-time. For DOF simulation, a depth buffer (also called z-buffer sometimes) can be employed to store, update and fetch the depth of each pixel with respect to camera position. Approaches based on stereo matching also exist but they suffer from the ill-posed correspondence problem. Especially for natural scenes, the results include noise and artifacts due to depth discontinuities which influence the performance of real-time applications. In case of aerial or satellite imagery, digital elevation models (DEMs) are often available through land surveys, or triangulation procedures using a combination of sensors. A DEM, when available, can function similarly to a depth buffer when stereo matching may not be reliable in real-time. We employed DEMs to obtain a smooth rendering without visual artifacts, which is especially important for perceptual tests.

2.4. Combined Model and Implementation

Albeit few, existing user studies on individual perceptual models (e.g., CSF based foveation or color) offer promising results [DCM04, DBS*09, WWHW97]. A combined model, which would have the highest fidelity to the HVS, is rarely found in the literature and there are no empirical studies linked to such combined models to our knowledge. Below, we describe our implementation where combinations of various HVS models can be applied and spatial parameters can be adjusted, which, in turn, will serve as a testbed both for computational and perceptual experiments.

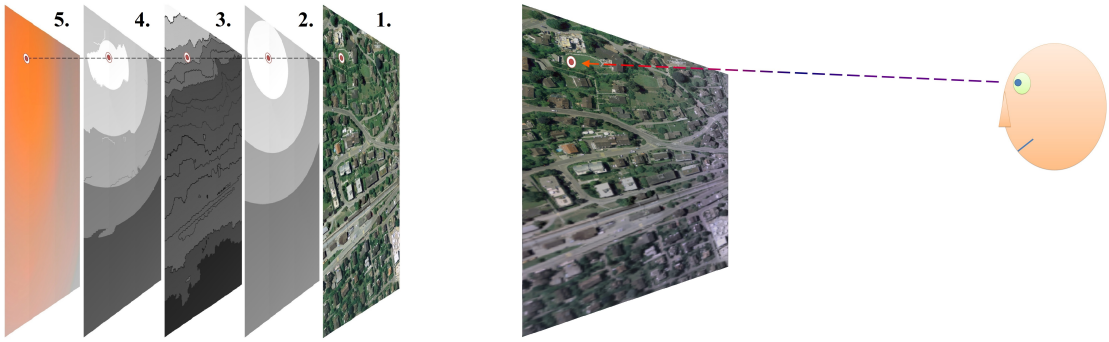

Figure 2: *The right-most image shows the output from the combined model when the viewer's point of interest (POI) is at the upper left corner of the image (marked by a red dot). Note that the image resolution is degraded and the color is fading out from the POI towards the periphery. Towards left, we can see 1) Uniform resolution input image ©2015 swisstopo (BA15026) 2) CSF based foveation 3) DEM 4) DOF simulation 5) Color degradation mask.*

For real-time gaze contingent rendering, we used OpenGL (`https://www.opengl.org/`) and implemented our models in fragment shaders (GLSL language). The system works with mouse as well as gaze input (through eye tracking). The design is similar to a Model-View-Controller (MVC) in which model and view has no direct communication. Model component encodes bmp, tiff, jpeg and shp files, correspondingly loads raster and vector graphics to buffers, and binds these buffers to texture samplers. In View component GLFW library (`http://www.glfw.org/`) is used to present the OpenGL context in a window and to maintain user interactions with keyboard and mouse callback functions. Controller part is responsible for the orchestration of interaction events and the application logic that is composed of:

- Initializing the window and VPM related parameters
- Rendering raster and vector layers in OpenGL main loop
- Fetching mouse or eye tracker input in mutual exclusion
- Simulating the POI based on pre-calculated pixel coordinates (for computational testing)
- Recording the scene in raw or compressed formats
- Generating task list for the factorial user study design.

In the final stage of rendering pipeline (rasterization in fragment shaders), a mipmap mechanism was employed for the gaze contingent LOD adjustment. For each pixel, a new LOD was calculated (mipmaped) with respect to a weighted-Euclidean-distance between this pixel and the current POI. For the foveation model, the weighting was decided based on the CSF. For the DOF simulation, we used a digital elevation model (DEM: the third image in Figure 2). For foveation and DOF simulation, image resolution is reduced by the exponents of 2 (i.e., mipmapping principle is used also for 3D space). This is illustrated in the second and fourth images respectively in Figure 2. The change in resolution is illustrated with shades of gray starting with white (original resolution). For the color model, the weighting was based on the color degradation mask which was conveyed to the fragment shader as a second raster image. Color reduction is illustrated with the shades of orange (full color) in the fifth image in Figure 2. Our combined model first employs the CSF and reduces the resolution in 2D space, then employs the DOF model to further reduce resolution in 3D space. If you compare the second and fourth images in Figure 2 you can see the additional resolution degradation coming from DOF. Finally, the color model is used to adjust the chromaticity of each pixel.

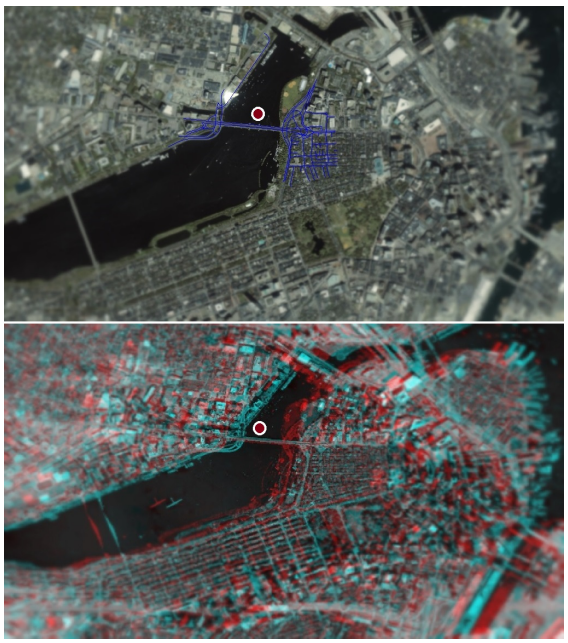

Figure 3: *Foveation with vector culling (top) and stereo foveation (bottom). Viewer's gaze (POI) is at the red dot outlined with a white line in all cases. Vector data: Office of Geographic Information (MassGIS). Raster data: Image ©2015 DigitalGlobe, Inc. Anaglyph image: Stereo GE Browser Version 0.14 ©Masuji SUTO & David Sykes 2009.*

The testbed provides an option to georeference images, thus, similarly to geographic information systems software, overlaying vector and raster data sets in the same visualization environment is possible. The vector overlay feature was further developed to provide "culling" of the vector layer from the peripheral regions (Figure 3 top). While the usefulness of this feature remains to be tested, it demonstrates that the HVS-inspired approaches used for modifying vector and raster data sets can be combined for future testing. Additionally, all testbed features can be viewed stereoscopically. This feature was implemented because satellite and aerial imagery are often (but not always) viewed stereoscopically in geographic studies, providing a strong depth cue in the presence of only few other depth cues in orthogonal aerial views (Figure 3 bottom).

3. Results and Discussion

3.1. Computational Experiments

In initial computational experiments, we measured *the amount of data* that was necessary to generate the images that are processed with VPMs. Therefore, we used image compression to study how much data would be discarded frame-to-frame if the uniform resolution images were visualized in a gaze-contingent manner (with simulated POIs). In particular, we analyzed the change of image compression ratio (i.e., uncompressed image size divided by compressed image size) as a function of the implemented VPMs. Using each model in the testbed, we generated a series of raw images as we altered the POI on a regular grid to balance for the LOD changes across the entire display. Following this, we compressed the generated images with the standard JPEG codec. We documented how the compression ratio was affected before and after the uniform resolution images were processed with the proposed VPMs. For instance, as the POI was closer to the image periphery, a larger part of the image remained out of focus (more blurred) thus we expected to observe a higher level of data reduction. Additionally, we analyzed the relationship between the compassion ratio and the image resolution. In the experimental setup for computational validation tests (desktop setup), the viewing distance was adjusted to 53 cm. Maximum LOD was 4 (see in second image in Figure 2) and the first level (white area where the image resolution is preserved) covers about 15° eccentricity. The size of the color mask (see in fifth image in Figure 2) was adjusted to 20° as it was suggested in [DBS*09] for simulating the normal viewing conditions. Depending on the POI location, the initial results indicate that our combined model leads to 2 to 10 times better compression ratios than the uniform resolution images when both are processed through the standard JPEG compression. The results also showed that VPMs decrease the average entropy by employing lower resolution levels of the mipmap and introducing more blur to the out of focus regions. The reduction in compression ratio was proportional to the image resolution. Images that were processed with the VPMs require less memory space than the uniform resolution counterparts, provided that all images are saved in a compressed format.

3.2. User Testing

In terms of user experience, in a controlled lab study, we measured participants' efficiency (task completion time in seconds) and effectiveness (response accuracy) as a function of the implemented models (i.e., independent variable) in comparison to uniform images. Our working hypothesis reads as follows: *Removing perceptually irrelevant data (i.e., using gaze contingent visualizations) does not affect the user experience*. 12 participants were asked to perform visual search tasks on series of aerial images (see an example on the image 1 in Figure 2) in 60-minute sessions. We used a large rear projection screen, thus the viewing distance was 3,5m and the display covered about 65° of participants' visual field of view. Our non-contact remote eye tracker allowed participants to sit and act naturally. The VPMs and the aerial images were presented in a Latin squares randomization to avoids learning effect. The location of the target (a circular map symbol) was also systematically changed to avoid bias from its position. The results from this study showed that about 70% of the participants did not notice any visual artifacts and efficiency and effectiveness were not harmed by the GCDs. In other words, participants' performance with visually degraded imagery was no different than with the uniform imagery (one-way repeated measures ANOVA: $p > .05$). This finding, combined with the fact that 70% reported that they did not notice any difference is reassuring in terms of "perceptually lossless" suggestions.

3.3. Summary and Outlook

Our implementation enables us to study the computational and human factors in using VPMs in various combinations. Can we improve the user experience as we discard the content deemed "perceptually irrelevant"? Our experiments demonstrate that our testbed is a useful tool in gaining new knowledge on computational and user performance with the GCD paradigms. At this point, both the computational experiments and the user study results indicate that the HVS-based data reduction solutions are competitive, and encourages further research. Further computational testing and a larger user study has already been conducted and will be published in near future.

References

[BB05] BARTHOLOMÉ E., BELWARD A. S.: GLC2000: a new approach to global land cover mapping from Earth observation data. *International Journal of Remote Sensing 26*, 9 (May 2005), 1959–1977. 1

[Bc11] BEKTAŞ K., ÇÖLTEKIN A.: An Approach to Modeling Spatial Perception for Geovisualization. *Procedia - Social and Behavioral Sciences 21* (Jan. 2011), 53–62. 2

[Bc14] BERNABÉ POVEDA M. A., ÇÖLTEKIN A.: Prevalence of the terrain reversal effect in satellite imagery. *International Journal of Digital Earth* (2014), 1–24. 1

[BÇS12] BEKTAŞ K., ÇÖLTEKIN A., STRAUMANN R. K.: Survey of true 3d and raster level of detail support in gis software. In *True-3D in Cartography*. Springer Berlin Heidelberg, 2012, pp. 43–65. 1

[BGS00] BARCLAY T., GRAY J., SLUTZ D.: Microsoft TerraServer. *ACM SIGMOD Record 29*, 2 (June 2000), 307–318. 1

[BRB06] BACIC I. L. Z., ROSSITER D. G., BREGT A. K.: Using spatial information to improve collective understanding of shared environmental problems at watershed level. *Landscape and Urban Planning 77*, 1-2 (June 2006), 54–66. 1

[Ç09] ÇÖLTEKIN A.: Space-variant Image Coding for Stereoscopic Media. In *2009 Picture Coding Symposium* (May 2009), IEEE, pp. 1–4. 2

[cDBV14] ÇÖLTEKIN A., DEMSAR U., BRYCHTOVA A., VANDROL J.: Eye-hand coordination during visual search on geographic displays. In *Proceedings of the 2nd International Workshop on Eye Tracking for Spatial Research, GIScience2014* (Vienna, Austria, 2014). 1

[cR11] ÇÖLTEKIN A., REICHENBACHER T.: High Quality Geographic Services and Bandwidth Limitations. *Future Internet 3*, 4 (Dec. 2011), 379–396. 1

[DBS*09] DUCHOWSKI A. T., BATE D., STRINGFELLOW P., THAKUR K., MELLOY B. J., GRAMOPADHYE A. K.: On spatiochromatic visual sensitivity and peripheral color LOD management. *ACM Transactions on Applied Perception 6*, 2 (Feb. 2009), 1–18. 2, 4

[Dc07] DUCHOWSKI A. T., ÇÖLTEKIN A.: Foveated gaze-contingent displays for peripheral LOD management, 3D visualization, and stereo imaging. *ACM Transactions on Multimedia Computing, Communications, and Applications 3*, 4 (Dec. 2007), 1–18. 2

[DCM04] DUCHOWSKI A. T., COURNIA N., MURPHY H.: Gaze-contingent displays: a review. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society 7*, 6 (Dec. 2004), 621–34. 2

[DHG*14] DUCHOWSKI A. T., HOUSE D. H., GESTRING J., WANG R. I., KREJTZ K., KREJTZ I., MANTIUK R., BAZYLUK B.: Reducing Visual Discomfort of 3D Stereoscopic Displays with Gaze-contingent Depth-of-field. In *Proceedings of the ACM Symposium on Applied Perception* (New York, NY, 2014), SAP '14, ACM, pp. 39–46. 2

[FCD*14] FUSTES D., CANTORNA D., DAFONTE C., ARCAY B., IGLESIAS A., MANTEIGA M.: A cloud-integrated web platform for marine monitoring using GIS and remote sensing. Application to oil spill detection through SAR images. *Future Generation Computer Systems 34* (May 2014), 155–160. 1

[GOT*11] GIRI C., OCHIENG E., TIESZEN L. L., ZHU Z., SINGH A., LOVELAND T., MASEK J., DUKE N.: Status and distribution of mangrove forests of the world using earth observation satellite data. *Global Ecology and Biogeography 20*, 1 (Jan. 2011), 154–159. 1

[GP98] GEISLER W. S., PERRY J. S.: A real-time foveated multiresolution system for low-bandwidth video communication. In *in Proc. SPIE* (1998), pp. 294–305. 2

[GSZ14] GUO W., SHE B., ZHU X.: Remote Sensing Image On-Demand Computing Schema for the China ZY-3 Satellite Private Cloud-Computing Platform. *Transactions in GIS 18* (Nov. 2014), 53–75. 1

[HM97] HAHN P., MATHEWS V.: Perceptually Lossless Image Compression. In *Proceedings DCC '97. Data Compression Conference* (1997), IEEE Comput. Soc. Press, p. 442. 2

[Lin04] LINDE I. V. D.: Multi-resolution image compression using image foveation and simulated depth of field for stereoscopic displays. In *Proceedings of the SPIE* (2004), vol. 44. 2

[LWC*02] LUEBKE D., WATSON B., COHEN J. D., REDDY M., VARSHNEY A.: *Level of Detail for 3D Graphics*. Elsevier Science Inc., New York, NY, USA, 2002. 1

[MB13] MARSHALL W., BOSHUIZEN C.: Planet Labs' Remote Sensing Satellite System. In *Small Satellite Conference* (2013). 1

[MM96] MANJUNATH B., MA W.: Browsing large satellite and aerial photographs. In *Proceedings of 3rd IEEE International Conference on Image Processing* (1996), vol. 1, IEEE, pp. 765–768. 1

[MWS13] MARDELL J., WITKOWSKI M., SPENCE R.: A Comparison of Image Inspection Modes for a Visual Search and Rescue Task. *Behaviour & Information Technology* (Aug. 2013), 1–30. 1

[Red01] REDDY M.: Perceptually optimized 3D graphics. *IEEE Computer Graphics and Applications 21*, 4 (2001), 68–75. 1, 2

[Rok96] ROKITA P.: Generating depth of-field effects in virtual reality applications. *IEEE Computer Graphics and Applications 16*, 2 (Mar. 1996), 18–21. 2

[SBA14] SCOTT G. J., BACKUS K., ANDERSON D. T.: A multilevel parallel and scalable single-host GPU cluster framework for large-scale geospatial data processing. In *2014 IEEE Geoscience and Remote Sensing Symposium* (July 2014), IEEE, pp. 2475–2478. 1

[SBSS14] STODLE D., BORCH N. L. T., SOLBØS. A., STORVOLD R.: High-performance visualisation of UAV sensor and image data with raster maps and topography in 3D. *International Journal of Image and Data Fusion 5*, 3 (May 2014), 244–262. 1

[SSGM04] SAITO K., SPENCE R. J. S., GOING C., MARKUS M.: Using High-Resolution Satellite Images for Post-Earthquake Building Damage Assessment: A Study Following the 26 January 2001 Gujarat Earthquake. *Earthquake Spectra 20*, 1 (Feb. 2004), 145–169. 1

[WWHW97] WATSON B., WALKER N., HODGES L. F., WORDEN A.: Managing level of detail through peripheral degradation: Effects on search performance with a head-mounted display. *ACM Trans. Comput.-Hum. Interact. 4*, 4 (Dec. 1997), 323–346. 2

[YW09] YANG B., WEIBEL R.: Editorial: Some thoughts on progressive transmission of spatial datasets in the web environment. *Computers & Geosciences 35*, 11 (2009), 2175–2176. 1

[YZP08] YANG Z., ZHU Y., PU Y.: Parallel Image Processing Based on CUDA. In *2008 International Conference on Computer Science and Software Engineering* (2008), IEEE, pp. 198–201. 1