An Approach to Modeling Spatial Perception for Geovisualization

Kenan Bektaş*a, Arzu Çöltekin*a

aDepartment of Geography, University of Zürich, Winterthurerstr. 190, Zürich 8057, Switzerland

Abstract

The amount and intensity of geographic information make it complicated to perform geographic data management and spatial analysis. In this paper we propose a novel approach exploring technical and perceptual issues in reducing information intensity; especially in the representation of geographic information. Foveation is a perceptually lossless degradation of the intensity of visual information. We apply established foveation algorithms on georeferenced raster and vector data sets and obtain visualizations in the form of foveated stereoscopic gaze contingent displays. Our concept is called as Geofoveation, and it is informed by vision research, computer graphics and geovisualization. In this paper, we discuss geofoveation as a means to demonstrate how spatial perception can be modeled and integrated into geographic visualizations. We hope to contribute to the discourse on intelligent geographic data management as well as provide initial steps to knowledge integration from certain aspects of spatial perception to geovisualization.

Keywords: Foveation; geovisualization; gaze contingent displays; data management; spatial perception; geofoveation;

1. Introduction

As a result of technological innovations, the amount of the geographic data has continuously been growing. Today, the availability and intensity of the information for a given location on earth is higher than ever before and high resolution geographic data can add to this information intensity dramatically. For the analysis of geo-spatial problems in this information-rich setting, the general properties of the geographic information have to be considered [14] and problem specific issues should be studied with critical spatial reasoning accordingly [15]. Spatial reasoning is a form of spatial thinking and spatial thinking, as a dynamic process, allows us to describe the structure of spatial objects and to explain and predict the interactions between these objects in a real or imaginary world [6]. An individual performing spatial thinking typically generates hypotheses, makes predictions and explanations for identifying patterns in an environment [33]. However, one should keep in mind that there is some degree of uncertainty in conception, measurement, representation and analysis of geographic data [21]. The information intensity and uncertainty brings some difficulties for an individual facing a geo-spatial
analysis task [9, 32]. How can we filter and organize the information, so that the geo-spatial analyst can evaluate it in an effective and efficient manner? We may find an answer to this question potentially by studying the nature of our spatial perception.

The human brain and the eyes do not process all of the available visual information in a uniform fashion as presented on a scene (i.e. real, augmented or virtual). There are bottom-up and top-down approaches explaining the working mechanisms of this non-uniform spatial vision, taking several perceptual (bottom-up) and cognitive (top-down) factors into account. First and foremost, the bio-optical configuration of the eye determines the limits of our visual perception. Additionally, cognitive factors, such as subjective preferences, needs and level of concentration may affect what we understand from what we see. In the frame of this paper, we will focus on the perceptual processes.

Eye tracking is one of the very well-suited methods for studying spatial perception and has been used for decades (e.g. [38]). Eye tracking technology can be used for studying scene perception as well as for the validation of a geovisualization design. For example, based on the structure of visual information and biological mechanisms, Swienty (2008) studied the subjective relevance of the geographic information in satellite imagery and vector maps and used eye tracking technology for the validation of his work [35]. Hiroyuki (1997), on the other hand, investigated the effect of gaze contingent visualizations on the perception of peripheral information using eye tracking [16].

Gaze-contingency, as a paradigm, is an interesting visual information filtering approach, and when coupled with eye tracking, it can be an elegant addition to the efforts of reducing the perceptually irrelevant information [8]. Gaze-contingency paradigm uses models of human spatial perception, and can be applied to geographic data representations. This paradigm can potentially be used for Geographic Information Systems (GISs) and Geo-Virtual Environments (GeoVEs). GIS and GeoVEs are used to study complex and multidimensional geographical phenomena and to interpret related dense datasets by the help of 2D and 3D visualizations as well as virtual reality (VR) systems [10]. Inspired from [34], we summarize and categorize the information intensity related issues in GIS and GeoVEs as follows:

- Technical: How do we change the amount of detail for different scale levels?
- Perceptual: How much visual realism is enough for a given task?
- Cognitive: How much detail is necessary for the map to be understood by its intended audience?
- Experimental: How do we change the amount of detail as the purpose of the visualization changes?

This paper positions itself in the agenda above at the intersection of technical and perceptual considerations. We deal with information intensity based on specific aspects and knowledge on human vision and spatial perception. We present an approach that provides a perceptually lossless compression for geo-referenced graphics and visualizes these graphics in a stereoscopic gaze contingent display. Our approach does not answer the questions listed above, but instead, explores the possibility and efficiency of integrating some mathematical models of vision and spatial perception into geovisualization.

2. Conceptual Development

2.1. Visual Perception

Seeing, or visual perception, is one of the five senses that humans have and it plays an important role in developing geo-spatial skills such as navigation and space interpretation as well as tasks such as food search and using tools [37]. Relevant to our study, Ware (2004) defines one strategy for visualization design; i.e., transforming the data so that it appears like a familiar environment; a kind of ‘data landscape’. It is often hypothesized that such a ‘landscape’ would potentially stimulate the visual attention as in the natural world and therefore would be easy to process for the viewers [29]. Visual attention is a selective gating mechanism which enhances visual processing in the presented environment, so that the observer deals with certain stimuli more effectively than others [30]. Visualization designs that take psychophysical limits of human vision into account can be useful at tasks that require visual analytics. Visual
analytics is defined as an automated analytical reasoning process combined with interactive visualizations to generate knowledge from large amounts of data and it is widely explored in GIScience in recent years [18].

2.2. Graphic Data and LOD Management

Many data types that are fundamental to geographic processes and analyses (e.g. terrain data, city models) are typically generated and stored in high resolution. This high spatial resolution provides a good visual fidelity, but it requires more computational processing. Therefore, it is necessary to find a balance between technical (e.g. processing power, storage, bandwidth limitations) and perceptual (e.g. visual fidelity, visual quality, visual clutter) issues. In computer graphics, one approach to address this tradeoff is level of detail (LOD) management. If computational resources are sufficient, highly detailed datasets can be stored and visualized, however, most of the time the datasets are prohibitively large for this and human visual information processing capacity is limited. That is, when the complexity of the data hinders real-time rendering or the visualization designer (e.g. a cartographer) decides to remove redundant details, then less detailed versions are preferred. Alternatively, highly detailed and less detailed representations can be combined and visualized on the same display.

LOD management requires certain decisions which are based on object’s location and its geometrical properties as well as the viewer’s location and visual capabilities [23]. In view independent LOD management, multiple versions of graphical objects can be created in a preprocessing step based on their size or distance to the viewer. On the other hand, in view dependent LOD management, visualization is dynamically modified considering perception based criteria like eccentricity, contrast sensitivity (CS), velocity of the objects with respect to the user and (stereoscopic) depth of field (DOF). LOD management is commonly employed for transmission and visualization of terrain data as well as city models [36]. By applying efficient LOD management we obtain considerable compression, however this is not ‘traditionally’ listed as a compression approach; there are more established compression and spatial partitioning techniques (for a review see [24]). The conventional compression techniques and some LOD management approaches are already implemented for GISs, however, none of them consider the properties of human visual system (HVS) and none of them addresses the limitations and strengths of human spatial perception [2].

The HVS does not process visual information uniformly; therefore presenting uniform resolution visualizations with a single LOD is in fact a waste of resources [27]. Variable resolution displays (VRDs), on the other hand, have more than one LOD on a single representation. They work with the principle of providing high-resolution information on the point/s of interest (POI) and reduced resolution on the remaining part of the display. Gaze contingent displays (GCDs), a type of VRDs, use eye-tracking equipment to identify the POI (where people look is presumably where they attend). GCDs provide computational efficiency as they discard a large chunk of information in the peripheral regions and may reduce the simulator sickness problems if implemented for stereoscopic displays [7, 20, 27]. Similar to the human visual perception of the natural world, a non-uniform resolution representation is visualized on a GCD. The perceptually lossless degradation of the image resolution from the POI towards periphery is known as foveation in computer vision literature [4]. However, it is important to note that during visual perception the brain does not only process information from the point where the eyes are converging or focusing but from a larger area called useful field of view or functional field of view [25].

While not exclusively ‘geo’ visualization, GCDs present a relatively novel and experimental approach for visualization. For example, GCDs can be an alternative approach to widely used adjacent displays (multiple displays) and chessboard displays (multiple windows on single displays) [26]. There are GCDs developed for 2D, 3D and stereoscopic visualizations [12]. Duchowski (2004) reports that, some early implementations have suffered from technical difficulties and high costs of necessary equipment [12].
However, current performance of the software and cost of hardware are not necessarily a bottleneck for the application of GCDs to geographic data processing.

During gaze contingent visualization, multi-resolution images are sampled from a hierarchical data structure. This operation can be considered a form of image fusion. Image fusion techniques are widely used for medical imaging as well as remote sensing and geovisualization applications. The output of an image fusion operation is the composition of some input images from the same region of interest. However, in practical applications the available input may have either same or different data structures (e.g. raster and/or vector) or they may be produced by different sensors (e.g. satellite imagery from different spectral intervals). Ideally, if the fusion is performed adequately, the observer does not recognize the transitions between different layers having different resolutions or structures.

2.3. Modeling Spatial Perception for Geovisualization: Geofoveation

As mentioned in the introduction section, there are many reasons for reducing and filtering geographic data. Our motivation for creating geofoveation, an extended application of foveation algorithms to geospatial graphics, stems from the needs that occur in sub-domains of geography. For example, remote sensing imagery is typically very large in size and resolution; resulting in very large georeferenced image databases. In some cases, further data is provided by sensors other than remote sensing satellites which provide complementary information [25]. For specific analyses it may be necessary to fuse these as a single map. In such cases, foveated gaze contingent displays may provide computational efficiency and more control over the spatial analysis.

Geographic data is typically composed of multiple layers. Vector overlaid raster maps and textured vector data are commonly used in interactive map services and GeoVEs. For our study we use a multilayer (i.e. raster and vector) georeferenced data. We analyze the potential computational efficiency obtained with foveation methods for the visualization of high resolution data, and spatial relations between different layers before and after the application of foveation. Additionally, foveated visualization techniques can be considered similar to focus+context displays and interactive highlighting techniques [19, 31], which are used with the aim of providing an effective visual analytics tool for geo-spatial data.

Our understanding of space depends on our perception of space; therefore we study the relation between the HVS and spatial thinking. LOD management techniques and perceptual analysis have been extensively studied for 2D visualizations and the research and innovations in this field are ongoing. However, it is still necessary to study the effect of depth perception on the visual attention and especially on the perception based 3D visualization [23, 30]. Geofoveation approach includes modeling three-dimensional spatial perception and its application to stereoscopic 3D geovisualization. Overall, geofoveation makes use of the gaze contingency paradigm, stereoscopic visualization and foveation as a LOD management technique for geo-referenced graphics [1].

In the conceptual development of a spatial perception model for geovisualization, we define the properties of an adequate fusion operation that is optimized for geographic data as follows:

- In a two-dimensional non-uniform visualization, the resolution should degrade towards the periphery in proportion to the structure of human fovea, i.e. the pixel size or count should correspond to the photoreceptor density (con cells are dense in the center, much less in the periphery).
- Degree of contrast and color sensitivity of photoreceptors should be included in the model for radiometric considerations for an HVS inspired graphic representation.
- Eye movements should be detected in real time and the location of the gaze should be fed to the system as the POI.
- For stereoscopic gaze contingent visualization, POI should be identified in 3D space, ambient illumination and the relation between color and depth perception should be taken in to account.
- The fusion operation should be performed in real-time. Therefore, flicker effects and visual artifacts should be addressed for seamless visualization.
3. Methods

Relying on its conceptual and methodological framework, geofoveation project encapsulates and combines different approaches which are specialized for efficient processing and effective visualization of multi-resolution geo-spatial graphics data. The hierarchical and multi-resolution data structures are used for these purposes for decades. Here we will briefly present some of the approaches that we make use of in geofoveation project.

3.1. 2D Foveation for Image Pyramids

Although there are older publications, today a remarkable amount of work about multi-resolution image analysis relies on (or inspired by) the work of Burt and Adelson [3]. They present an image encoding technique which is based on down-sampling, weighting and finally up-sampling processes. Due to the resolution reduction in the first step, resulting hierarchical data structure is generally called as image pyramid. Space variant displays often use an image pyramid.

The purpose of developing space variant displays was to study (dis)functionality of the HVS as well as to provide an engineering solution to bandwidth related problems. For example, a space variant imaging system (SVIS)† presented by Geisler and Perry [13] basically makes use of foveation concept and implements image pyramids for the multi-resolution image coding with the purpose of bandwidth reduction [13]. The SVIS, as a standalone application, is able to locally perform image coding operations for static imagery and video. In a later publication, Perry and Geisler have successfully applied foveation methodology on real-time gaze contingent processing of colorful 640x304 video sequences, and achieved a processing performance of 28 frames per second [28].

As briefly explained earlier, for a system like SVIS, it is necessary to define a seamless resolution change across the foveated image to ensure image quality. Both in vision research and digital image processing literature, image quality analysis is frequently linked to the contrast sensitivity of the human eye. The term contrast sensitivity expresses how much the human eyes respond to different levels of luminance. Thus, contrast sensitivity function (CSF) is a measure of low-level perceptibility of visual stimuli that tests fidelity/performance tradeoff of simplifications. Geisler and Perry [13] use the CSF in equation (1) where, \( f \) is the spatial frequency in cycles per degree, \( e \) is the retinal eccentricity in degrees, \( CT_0 \) is the minimum contrast threshold, \( \alpha \) is the spatial frequency decay constant and \( \epsilon_2 \) is the half-resolution constant [13].

\[
CT(f, e) = CT_0 \exp \left( \alpha f \frac{e + \epsilon_2}{\epsilon_2} \right)
\]  

(1)

---

† Space Variant Imaging System: http://svi.cps.utexas.edu/
During image encoding, the individual portions of the levels from foveated pyramid have to be recalled and blended seamlessly. Perry and Geisler [28] explain the blending function that is used in SVIS in [28]. Based on the so called Goldmann perimeter they estimate the relative visual acuity of a healthy human eye, (see Fig 1.). The levels of the image pyramid, \( R_i \) have a fixed individual resolution (let’s say \( R_0 = 2R_1 \)). In [28] Perry and Geisler define: a transfer function \( T(R(x,y)) \), showing the level of amplitude attenuation in terms of spatial frequency, a blending function, \( B(x,y) \), for a seamless transition between adjacent levels (see equation 2) and the output image \( O(x,y) \) (see equation 3 and Fig 3.).

\[
B_i(x,y) = \begin{cases} 
1 & (x,y) \geq R_{i-1} \\
0.5 - \frac{T_i(R(x,y))}{T_{i-1}(R(x,y))} & R_i < R(x,y) < R_{i-1} \\
0 & R(x,y) < R_i 
\end{cases}
\]  
(2)

\[
O(x,y) = B(x,y)I_{i-1}(x,y) + (1 - B_i(x,y))I_i(x,y)
\]  
(3)

### 3.2. Wavelet Foveation

Image processing applications are typical examples of digital signal processing which can be investigated using the general rules of linear systems theory. The most common ways of analyzing and synthesizing digital signals are fourier (older) and wavelet (new) transformations. The basic principle behind both of these methods is to chunk a long and complex digital signal into manageable shorter pieces. In other words, the input signal is represented in terms of some orthogonal (i.e. linearly independent) basis vectors and coefficients. In Fourier analysis the basis vectors are cosine or sine functions but wavelets do not only have one group of basis. This makes the wavelets more flexible, thus they can be applied on various problem domains.

Chang (1998) defines the foveation \( (T_f)(x) \), by a smoothing function \( g: \mathbb{R} \to \mathbb{R} \), and a weight function \( w: \mathbb{R} \to \mathbb{R}_{\geq 0} \) [5]. It is represented with an integral operator applied on the original image \( f(t) \) as follows:

\[
(T_f)(x) = \int_{-\infty}^{\infty} f(t) \frac{1}{w(x)} g\left(\frac{t-x}{w(x)}\right) dt
\]  
(4)

Where \( w(x) = a|x - y| + \beta \), \( \alpha \) is the speed of resolution decay, \( \beta \) is the maximum resolution at \( y=t \), the point of fovea. Therefore, the foveation model in [5] is based on the Euclidean distance, which yields to a circular region of interest centered on the POI, \( y \). In 2D case the wavelet sub-bands are determined horizontally and vertically. They are denoted as LL (low-pass filtered rows and columns), HL (low-pass filtered to rows and high-pass filtered columns), LH (low-pass filtered to columns and high-pass filtered rows), and HH (low-pass filtered rows and columns). The multi-resolution structure of the foveated wavelet sub-bands are shown in Fig 2.

### 3.3. Geofoveation

In this study, we implement parts of the Geofoveation testbed; that is, foveation algorithms developed for raster images [5,13], and optimize them to study various geovisualization scenarios including 3D representations with multilayer raster and vector type georeferenced data.
An example scenario would be as follows: A test person sits in front of a GCD and the system foveates the field of view according to his eye movements, i.e. POI. The focus point may fall on to a part of or on the whole body of a specific geographic object such as a building or a section of a road or a river. The system can make use of GIS (for retrieving the object information) as well as eye tracking (as a proxy for the person’s visual attention) and can make an ‘educated guess’ to identify user’s interest. Based on these inputs, differently from other foveation approaches, geofoveation system can visualize the remaining
portion of the geo-object in high resolution even if it may fall in a region which should be foveated. This is an adaptive response which modifies the geovisualization design according to the user’s interest.

The final geofoveation testbed will provide an environment for obtaining perceptually lossless compression and a visual perception based manipulation of the georeferenced data which would lead to geo-foveated 2D, 3D and stereoscopic 3D visualizations. These sets of visualizations are an expected outcome of the testbed where the computational performance of different foveation methods will be compared for geographic tasks. The future testbed can also be used to study the relationship between foveated displays and spatial thinking and to answer a range of cognitive and experimental questions.

For maintenance and comparative testing purposes geofoveation test-bed is to be composed of several modules, namely; Eye Tracking, 2D Foveation, 3D Foveation, Motion Perception and Combined Spatial Vision. Currently, geofoveation testbed is able to foveate tif and jpg type geo-referenced 2D images. The vision based parameters are embedded into 2D foveation module but we plan to separate this in the future. Eye tracking module is under development so in current implementation POI is determined by the position of mouse cursor. We overlay this geofoveated raster data with georeferenced vector data that is extracted from a shape (.shp) file by using a region of interest centered at POI, see Fig 3.

4. Conclusions and Future Work

The fundamentals of our study are mainly based vision research, computer graphics and geovisualization. Therefore we have explained data, analysis and visualization related issues in geovisualization; some properties of human visual perception; basics of level of detail management, image fusion and gaze contingent visualization for computationally efficient visualization and effective analysis of geo-spatial data. We have briefly explained image pyramid based and wavelet based foveation algorithms for 2D raster images. We have also stated that the conventional compression techniques are widely used in current GISs but they do not take the properties of HVS into account [2], and presumably they are not related to spatial perception.

Geofoveation concept is a novel, interdisciplinary and experimental approach. Further modules will be developed and tested as the implementation progresses.

A next step is to apply standard mesh simplification algorithms [11] and view dependent vector simplification approaches [22] on to georeferenced vector data with an HVS-component. Following that, based on literature review in stereoscopic visualization, gaze contingent displays and perimetry research fields, we will determine an optimized functional visual field. Our current implementation is based on eccentricity and contrast sensitivity. To arrive at a more complete spatial perception model, we will include depth of field considerations, motion perception, color discrimination and ambient illumination in our near-future model. Also, we intend to further test and improve current stereoscopic foveation approaches (e.g. [7, 17]) to develop a foveated stereoscopic gaze contingent display for georeferenced data.

Current foveation techniques do consider eccentricity and contrast sensitivity of the human eye and they use experimented and established approximations for representing the foveated resolution change towards the peripheral region [5, 13] In these techniques, the geometric shape of the visual field which is used as a model for the resolution change is either rectangular or circular. However, the shape of the human visual field is not rectangular, neither it is circular. Applying a curve fitting algorithm on to an up-to-date perimetry data we will replace the rectangular and circular models with a more accurate model of visual field. We hypothesize that our model will provide a more effective and seamless visualization, because it will match better to the natural vision.

In this paper we presented the early steps, primary considerations and the general framework of the Geofoveation project. The technical details of our visual model (with perception parameters) will be presented in a follow-up publication.
Acknowledgement

This research is funded by Swiss National Science Foundation (SNF award number 200021_120434/1). We would like to thank GeoEye (www.geoeye.com) and Office of Geographic Information (MassGIS), Commonwealth of Massachusetts Information Technology Division for their cooperation and for providing the image and the shape file presented in Fig. 3.

References


