# A VISUALIZATION METHOD BASED ON FOVEATION

Arzu Çöltekin (Coltekin)

Department of Surveying Helsinki University of Technology arzu.coltekin@hut.fi

## ABSTRACT

A human visual system inspired process called *foveation* is an approach in managing the level of detail by controlling the region of interest. The scene is segmented according to the principles of our visual perception, modeling how the fovea behaves.

Foveation is interesting for designing robot eyes, and in bandwidth/resource hungry visualization tasks. It is also relevant for limited hardware capacity and processing very large datasets. Photogrammetry, Remote Sensing and Cartography deals with large images and scenes. Even though there are a few different approaches for managing the level of detail; foveation is not widely known or utilized in these fields.

This paper presents an introduction to foveation, an argumentation of how and why it could be useful for Geoinformation applications. An implementation of foveation for stereo imaging, called *Foveaglyph*, is also presented to demonstrate the concept.

## **1. INTRODUCTION**

### 1.1. Motivation

All subfields of geomatics deal with large amounts of data. It is seldom that a geo-dataset is conveniently small. Tterrain models are typically massive (Reddy et al., 1999).GIS deals with combined graphic and non graphic database, photogrammetry and remote sensing deal with high resolution images. The recent development on acquiring point clouds with laser scanning techniques also yield very dense and large datasets.

In computer graphics, "level of detail management" is for large visual datasets. The idea is developed over the thinking that "if we do not need to render an object, we should not render it and save the resources". This has been developed and applied over years by techniques that determined visibility of the model. Earlier considerations of "what is visible" were mostly based on geometric facts; view frustum was calculated, objects outside of it was not rendered, occlusions were taken into account etc. The development of level of detail concepts brought factors relating to human perception in calculating what is visible.

Human perceptions, particular to this research the visual perception, use highly efficient level of detail management. When the eyes locate their object of interest, they focus (accommodate) on the object and the rest of the scene is falls out of focus. It happens so that the resolution *slowly* decreases, therefore unless we pay special attention, we do not realize that the peripheral vision is not sharp.

This has inspired the development of new methods to control levels of detail in computer graphics and image processing. Applying further simplification to the scene to gain more resources if the user is not going to notice the difference and if it will not harm the operations makes perfect sense. It is attractive particularly for network applications and maps are most useful when they are mobile. Making digital maps for mobile devices means sending them over a network, or using them on devices not as strong as the regular office computers.

Considering the conflict between the amount of data and the need for mobility, it is clear that research towards making use of geo-data easier is useful. A through understanding and modelling of human vision is also relevant to close range photogrammetry research and development. This paper was written with a motivation to investigate the potential of a level of detail management technique called foveation for all fields of geomatics.

XXII International Cartographic Conference (ICC2005)

A Coruña, Spain, 11-16 July 2005 Hosted by: The International Cartographic Association (ICA-ACI) Produced by Global Congressos

ISBN: 0-958-46093-0

## 2. BACKGROUND

#### 2.1. Level of Detail Management for Geoinformation

Level of detail methods, Mesh simplification, Foveation

Majority of the maps are vector, and vector graphics are created based on point, line and polygons. The geometry of the smallest polygon is sometimes a triangle and "triangulation" is the process of creating a mesh to represent the terrain. Sometimes the basic shape is a quadrangle, and a grid is formed. In either case, the mesh is the common form to represent a continuous surface in vector graphics and it can have different resolutions. The denser the triangles, the more detail of the surface is represented.

**Mesh Simplification:** Though "more detail" is a desirable feature as clearly it would bear more information; it is not always wise to represent it all at once. The information should be kept in the storage, but should be presented only when it is needed and relevant. This logic have driven earlier "generalization" concept for Cartography for specific maps that addressed specific purposes. Today, it drives several researchers into a concept called "mesh simplification".

Hoppe defines mesh simplification as follows:

The meshes created by modeling and scanning systems are seldom optimized for rendering efficiency, and can frequently be replaced by nearly indistinguishable approximations with far fewer faces. At present, this process often requires significant user intervention. Mesh simplification tools can hope to automate this painstaking task, and permit the porting of a single model to platforms of varying performance. (Hoppe, 1996)

Hoppe's rather effective *progressive mesh* representation defines for an arbitrary triangle mesh a sequence of approximating meshes optimized for view-independent LOD. In his later work, Hoppe defines a framework for selectively refining an arbitrary progressive mesh according to changing view parameters (Hoppe 1998).

View dependent LOD can be determined based on a number of criteria.

**LOD Selection Criteria:** The following are listed as LOD selection factors (compiled from Reddy 1997, Constantinescu, 2001, Luebke et al 2003):

- Distance
- Size
- Priority
- Hysteresis
- Environments Conditions
  - Perceptual factors
    - Eccentricity
    - Velocity
    - Depth of field

Size LOD and Distance LOD also rely on human perception, therefore can be –and sometimes areconsidered under perceptual factors. However, they are the most common LOD applications and they often appear separately.

In this paper, we focus on two of the perceptual factors, which cover the visual perception of the static scenes; the eccentricity LOD (2D foveation) and the depth of field simulation; the combination of the two amounts to 3D foveation. The implementation was done with stereoscopic displays in mind therefore it is more proper to call it stereo foveation. The concepts covered will hold relevant for all kinds of 3D datasets.

**2D versus 3D:** The traditional maps are two-dimensional. The need to represent the height has usually been fulfilled by using contour lines. Maps with contour lines are sometimes referred to as 2.5D maps. In digital media, 3D representations are a lot more useful than in the printed media as its visualization can be rather flexible, therefore we see more 3D terrain models, city models, architectural models and such. Because of the fact that now digital media allows better and easier 3D representations, 3D maps

have become attractive for purposes other than navigation and thematic mapping and cartography as well. The research presented in this paper deals with stereo 3D image visualization.

Games utilize 3D maps heavily, and other branches of entertainment, such as animated films also use techniques that are used for map making earlier. Same applies for almost all Virtual Reality (VR) applications. Thus, it is true that "Geographic Information Systems (GIS) and scientific visualization tools have begun to expand into each other's domains" (Ryhne, 1997).

#### 2.2. Foveation

Foveation is a scientific visualization method based on human perception. The term LOD is used mainly by the computer graphics community and the term foveation by the computer vision and image processing. Even though the link between the two is quite obvious, the literature on the two does not refer one another very much. We view foveation as a LOD management technique.

The main purpose of foveation is to provide compression to help the performance for the storage, computation and transfer of the large visual datasets. These visual datasets can be images, videos or 3D models. Other than providing compression, foveation is perceived as a smart Level of Detail management system for such tasks.

The smartness of it lies in the fact that it conforms to the human visual system's principles, which is reported to bring an advantage over the otherwise uncomfortable side effects of stereo viewing.

Foveation is a biologically inspired computer vision method. The term comes from the word *fovea*, the part in the eye that controls the human spatial vision. It is applied mostly in image processing and in robot eyes (cameras). Foveation reduces the level of detail gradually at the areas where the human eyes process the perceived detail, therefore foveation is a compression method.

Foveation is also a *space variant* level of detail control system. *Space variant* means that the resolution of the image or the model varies throughout the spatial domain. This happens according to a pattern or a mathematical model. In fact the term space variant expresses our 3D level of detail control in this thesis quite well.



Figure 1: An illustration of foveation just to demonstrate its principle idea. 6 levels of detail are visible.

Foveated images have been exploited in computer vision, especially in the context of active vision. But they are also useful in visualization, although this aspect is less explored. (Chang et al. 1997).

Photogrammetry works with stereo vision to recover depth from images. It deals with 3D modeling and measurements. For most of its tasks, high resolution images are the main input into a photogrammetric system. Videos too, can be utilized. The term videogrammetry refers to this kind of videographic photogrammetry (Coltekin, 2005).

An implementation of foveation for stereoscopic visualization is presented below, which should also give an idea for its potential for other areas of Geo-visualization. Stereo is a special medium, but it is still a spatial visualization task and the principles therefore are comparable.

## **3. FOVEAGLYPH**

*Foveaglyph* is an implementation of foveation for a stereo pair and the stereo visualization method is anaglyph. It is programmed in C, currently running on Linux, though it can be compiled for other systems as well. It performs foveation for 2D and 3D. The stereo feature naturally requires image matching to be able to calculate the depth information. For the image matching and the disparity calculation, Foveaglyph uses "Depth Discontinuities by Pixel-to-Pixel Stereo" (p2p) published by Stan Birchfield and Carlo Tomasi (Birchfield et.al.1996).



Figure 2: Schematic description of Foveaglyph.

The program can be run from the command line, or using the graphical user interface.

```
Usage: foveaglyph [options] left image right image
-h, --help
                              this help message
                              write output to FILE (anaglyph)
 -o, --output=FILE
 -m, --max-disparity=MAX
                              maximum level of disparity (200)
-p, --point-of-interest=Y,X Set the point of intereset for foveation
 -l, --level=N
                              Construct N levels of image pyramid
 -c, --camera-constant=N
                              Camera constant
 -B, --base-distance=N
                              Base distance between stereo cameras (mm)
 -2, --two-dimensional
                              Foveate only in 2D, do not use disparity.
                              Foveate only in 3D (default),
 -3, --three-dimensional
 -4, --3d-disparity
                              Foveate only in 3D, only using disparity info,
 -r, --write-rcfile
                              Write calculated coordinates to output rc.txt.
-D, --debug-level
                              Debug level.
```

Figure 3: Foveaglyph's command line options.

For 2D foveation, *foveaglyph* needs one input image, and a point of interest (PoI) specified by user. If no PoI is given, the centre of the image is regarded as point of interest. There are other parametric features too but if user does not specify anything, program defaults are used as in choosing a PoI.

Both 2D and 3D foveation uses a distance function, taking the specified PoI into account. A co-centric circular geometry is planned for 2D and again co-centric rings of spheres for volumetric foveation when it is done in stereo. The two options for 3D foveation are for the cases when the camera setup is known or unknown.

An image pyramid is built simply by super-sampling the image into coarser resolutions. To reconstruct the foveated image, we decide from which element of the image pyramid each pixel must come from.

## 3.1. Results

We provide one 3D foveated image here, which is particularly exaggerated to demonstrate the effect. If it is viewed towards it periphery, and objects that are closer or farther away from the point of interest, it can be observed that the pixel sizes grow.



**Figure 4**: The 3D foveated analyph output from *foveaglyph*, with 5 levels of detail. The human eye will foveate with 35 levels of detail as reported in Nakayama 1990.

This effect may not look desirable at first, but according to the facts about how much humans perceive, in a wide field of view display, this will not be noticeable for the user.

Indeed there are several user studies that report this. For example, in a study Kortum and Geisler reported that subjects reported very little perceptual difference between foveated and uniform images (Kortum et al. 1996). Also Watson *et al.* report that user performance was not significantly affected by the degraded peripheral display and concluded that eccentricity LOD provides a useful optimization tool (Reddy 1997, Watson et al. 1995).

**Compression**: Foveation is an efficient compression method that has been tested for images and videos (i.e. see Chang 1997, Perry & Geisler, 2002). As for how much compression *foveaglyph* provides, that very much depends on the point of interest. In 2D foveation, if the PoI is closer to the periphery the compression gain will be bigger (more pixels come from the lower resolution images of the pyramid). In 3D foveation, the position of PoI is even more determining, because if the space

around the PoI is populated, there will be more information – so not only the location of the PoI, but also the content of the image has an effect on the results.

These said tests with 3 levels of detail and varying PoI's on a 630 x 480 image has given compressions between 61% and 89%. A more detailed account of this process is available in the upcoming PhD work of the author.

## 4. CONCLUSIONS

This has been a brief description of a complex phenomenon. The way the human visual system works is rather economical, though still occupies a good amount of our resources:

Vision is the most powerful of the senses, and it is by far the most neurologically demanding, with over 70% of all sensory receptors in the human nervous system dedicated to its functioning (Marieb, 2000). In a more general way of saying; "visual intelligence occupies almost half of your brains cortex" (Hoffman, 2000). To sum it up, 70% of all receptors, 40+% of cortex and 4 billion neurons are dedicated to vision and we can see much more than we can mentally image (Ware, 2003).

A through understanding and modelling of human visual system interests researchers from several fields such as artificial intelligence, robot vision, image processing, virtual reality and computer vision. Geomatics research has become almost all "computerized" in last decades and since then, there are overlapping research interests with the mentioned fields.

This research gives a brief overview of some techniques to manage large visual datasets in a similar way as the human vision works, represents an implementation to demonstrate that the compression gain for image data is plausible and suggests that this concept can be extended to all geo-datasets successfully.

Further research might be directed at testing the current implementation in network applications (such as an internet transmission) and testing the concept with vector graphics. As it is, the implementation yields results only meaningful for raster data.

#### REFERENCES

Birchfiled, S., Tomasi, C., 1996, "Stanford University Technical Report" STAN-CS-TR-96-1573, July 1996 – The code, the algorithms and related papers for "Depth Discontinuities by Pixel-to-Pixel Stereo" is published at: http://robotics.stanford.edu/~birch/p2p/, last visited in May 2005.

Chang, E.C., Yap, C., 1997, "A Wavelet Approach to Foveating Images", Proceedings of 13<sup>th</sup> ACM Symposium on Computational Geometry, pp. 397-399, 1997.

Coltekin, A., 2005, Stereo-foveation for anaglyph imaging, Proceedings of SPIE Volume: 5664 Stereoscopic Displays and Virtual Reality Systems XII, Publication Date: Mar 2005; 682 pages; 70 papers; Softcover; In print

Constantinescu, 2001, Levels of Detail: An Overview, First NTNU CSGSC, May 2001

Hoffman, D.D., Visual Intelligence, How We Create What We See, p.202, ISBN 0--393-31967-9 pbk. 2000, W.W. Norton & Company Ltd. New York, NY 10110

Hoppe, H. 1996. Progressive Meshes. Proceedings of ACM SIGGRAPH 1996, pages 99-108.

Hoppe, H., 1997, View-dependent refinement of progressive meshes. ACM SIGGRAPH 1997, 189-198.

Hoppe, H., 1998. Efficient Implementation of Progressive Meshes. In Computers & Graphics Vol.22, No.1

Kortum, P., Geisler, W. 1996. Implementation of a foveated image coding system for image bandwidth reduction. In Human Vision and Electronic Imaging, SPIE Proceedings, Vol. 2657, pp. 350-360, 1996.

Luebke, D., Reddy, M., Cohen, J.D., Varshney A., Watson, B., Huebner, R., 2003. Level of Detail for 3D Graphics. Textbook. Morgan Kaufmann, series in Computer Graphics and Geometric Modeling. ISBN 1-55860-838-9. URL: http://www.lodbook.com/.

Marieb, E.N., 2000. Essentials of Human Anatomy & Physiology. Benjamin/Cummings (imprint of Addison Wesley Longman), 6 edition.

Nakayama, K.,1990. Properties of early motion processing: Implications for the sensing of egomotion. In R.Warren and A.H.Wertheim, editors, The Perception and Control of Self Motion, pages 69--80. Lawrence Erlbaum, Hillsdale, NJ.

Perry J.S., Geisler W.S., 2002. Gaze-contingent real-time simulation of arbitrary visual fields. CA: Human Vision and Electronic Imaging, Proceedings of SPIE 2002, San Jose.

Reddy, M. 1997. Perceptually Modulated Level of Detail for Virtual Environments. Ph.D. Thesis. University of Edinburgh. Also can be found at URL: http://www.martinreddy.net/percept/ as of 11.4.2005

Reddy, M., Leclerc, Y. G., Iverson, L. and Bletter, N., 1999. "TerraVision II: Visualizing Massive Terrain Databases in VRML". *IEEE Computer Graphics and Applications (Special Issue on VRML)*, 19(2): 30-38.

Rhyne, T.M., 1997, "Going Virtual With Geographic Information and Scientific Visualization," *Computers and Geosciences*, Vol. 23, No. 4, 1997, pp. 489-491

Watson, B., Walker, N. and Hodges, L. F. (1995). A User Study Evaluating Level of Detail Degradation in the Periphery of Head-Mounted Displays, *Proceedings of the FIVE '95 Conference*, QMW, University of London, UK, pp. 203–212.

Ware, C., 2000, *Information Visualization: Perception for Design*. San Fancisco: Morgan Kaufmann. ISBN 1-55860-511-8