

# FOVEATION SUPPORT AND CURRENT PHOTOGRAMMETRIC SOFTWARE

A. Çöltekin

PL1200, Helsinki University of Technology  
02015, Espoo-Finland  
[Arzu.Coltekin@hut.fi](mailto:Arzu.Coltekin@hut.fi)

**Commission V, WG V/2**

**KEY WORDS:** Compression, Performance, Virtual Reality, Vision Sciences, Visualization, Foveation, Level of Detail, Photogrammetry

## **ABSTRACT:**

Foveation is a term used in computer vision to refer a technique of image compression. This type of compression takes the principals of human vision into account, trying to mimic the way the “fovea” works. “The human visual system implements an elegant compromise between the competing goals of maximizing field of view, maximizing spatial resolution, and minimizing neural resources: It encodes a large field of view using a retina with variable spatial resolution, and then when necessary, uses high-speed eye movements to direct the highest-resolution region of the retina (the fovea) at specific points in the visual scene.” (Perry & Geisler, 2002)

We can consider foveation as a method of level of detail (LOD) management, or under the bigger umbrella of area of interest management (AOIM).

Foveation has been employed both when the compression was needed for the system itself and when the image was to be sent over a network where the bandwidth availability has always been a limiting factor. Consequently, foveation is an efficient technique of managing photogrammetric visualization, particularly on 1:1 projections or over the network.

This paper presents an introduction of this important technique to the field of photogrammetry.

## **1. INTRODUCTION**

### **1.1 Motivation**

Close range photogrammetry and computer vision techniques have always had overlapping areas of research and methodology. This paper discusses one method that is commonly used in computer vision and image processing but seemingly not at all in photogrammetry, the method is called foveation.

Foveation is, in simplest words, a compression method. It is applied in images and videos that are either presented in big screens, or transmitted via networks. The first thing that makes foveation different than other compression methods is its cognitive quality. It takes the human vision into account.

In the areas where the close range photogrammetry is used, but particularly when the results are visualized in virtual reality rooms (caves) and panoramic big screens, this method can be an effective and suitable way to compress the presented image.

Motivation of this article comes from the realization that the current photogrammetric software do not have the support for this method. Therefore it tries to introduce foveation to the field of photogrammetry while giving an overview of the literature on the topic.

### **1.2 Overview**

The developments in the computer hardware have been tremendous over last decades. These developments have influenced all technical fields, including the 3D graphics and Virtual Reality (VR) research. Modeling based on microscopic or telescopic images, optical or thermal, active or passive sensor data has become possible. At the same time, as visualization proved useful and fascinating, this created wide interest.

Gaming, scientific visualization, education, documentation of cultural heritage, architectural walk-throughs, city and terrain models for planning, navigational or informational purposes have all blended in with the immersive VR experience.

Even though what is possible today should be viewed with great respect, it would be fair to accept that it is far from being fully immersive, and the full extent of what is available is not available without expensive hardware settings. A more moderate version of a Virtual Environment (VE) is available on the Internet offering some interactivity, and that comes with a set of difficulties.

The amount of data in such environments is big: constructing detailed vector geometry with polygons and triangles alone can amount to gigabytes, when rendered graphics and photo-textures are added, even more. On top of that, if the “virtual world” includes audio and video, model will be yet bigger.

Managing such a massive dataset is difficult even with the games that come in stand-alone CDs, because assuming the latest hardware in the user-end means narrowing the market down. It is self-evident that publishing it on a thin wire network, like the Internet, requires a lot more consideration.

This is one reason why researchers put effort in trying to find out functional methods to reduce the amount of data to be transmitted.

In this paper, we will look closely into one approach in making the transfer of visual data more feasible for a thin wire network. This approach is called “foveation”; it has been successfully utilized as a compression method for 2D images meantime it has a claim that it is a more “natural” way for humans to see things. Similar approaches, exploring or including the same idea, have been called “perceptually-driven” or “gaze-contingent” by different researchers.

## 2. HUMAN VISUAL SYSTEM AND COMPUTER GENERATED VISUALIZATION

“In order to modulate the LOD of an object based upon its perceptual content, one must first possess an appreciation for the way in which the human visual system is designed and how it is believed to function.” (Reddy 1997)

The term “foveation” comes from the word “fovea”, which is the name for a region inside the human eye. The description of the phenomenon was given in many sources, in two recent notes, it is as follows:

“The human visual system implements an elegant compromise between the competing goals of maximizing field of view, maximizing spatial resolution, and minimizing neural resources: It encodes a large field of view using a retina with variable spatial resolution, and then when necessary, uses high-speed eye movements to direct the highest-resolution region of the retina (the fovea) at specific points in the visual scene.” (Perry & Geisler, 2002)

“Resolution is much greater at the fovea in the center of vision than in the periphery. The foveal area subtends about two degrees of visual angle, i.e. a disc about two cm in diameter held at arms length. If we move ten degrees off to the side, the amount of fine detail that can be perceived declines by a factor of ten. It is also the case that color vision mostly takes place in the fovea. However, the periphery of vision is very important to our sense of orientation and to our sense of self motion.” (Ware, 2003)

As can be seen in the Figure 1, in the next column the visual acuity distribution is symmetrical towards the temple and towards the nose.

Figure 2 illustrates the foveation over an example image.

**Figure 2:** Exaggerated visual illustration of foveal vision. Here is a 6-fold (6 levels of detail) foveation from the point of gaze, or in other words, 6 levels of details. Keep in mind that this is a very rough representation, compared to the 35-fold that the human eye can achieve.

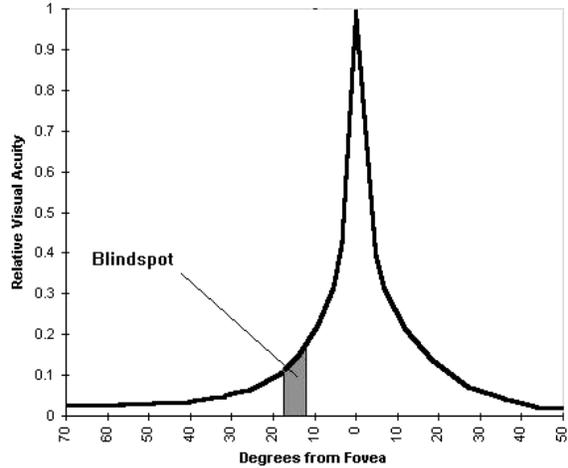


Figure 1: The human visual acuity distribution.

Normally it is stated that human eyes have roughly about a 35-fold difference (35 levels of detail) between the fovea and the periphery (Nakayama, 1990). The following figure shows an incremental use of blur function to reconstruct the scene employing foveation (Figure 2).



This special feature of human eyes was employed by the technical researchers as early as 1976. In his paper Clark suggested that objects can be simplified towards the periphery of the field of view. (Clark 1976)

For the Virtual Environments, this implies the reconstruction of the visual field according to what is best for human perception. Because [Virtual Environments are] “synthetic sensory experiences that communicate physical and abstract components to a human operator or participant. The synthetic experience is generated by a computer system that one day may present an interface to the human sensory system that is indistinguishable from the real physical world.” (Kalawsky, 1993)

Other than satisfying the human senses to provide immersion, this approach has another function, which is more relevant for publishing massive graphics and images on the Internet. As Murphy and Duchowski put it:

“The motivation behind such proposed gaze-contingent systems is to minimize overall display bandwidth requirements by reducing peripheral information in concordance with the perceptual limits of the Human Visual System (HVS). Efforts at providing peripherally degraded information date back to early eyeslaved flight simulators.” (Murphy & Duchowski, 2001)

### 3. TERMINOLOGY IN ACTIVE VISUALIZATION

As the reader might have noticed, the last quote did not mention “foveation” but a system that was “gaze-contingent”. As human vision is explored by different research groups, this concept is also called with different terms. The terms carry slightly different associations; for instance foveation, as a term, is used for systems with or without the eye-tracking while gaze-contingent term is associated with an eye-tracking system. They essentially mean the same thing.

For giving a name to this phenomenon, other than “foveated” and “gaze-contingent”, also “Eccentricity LOD” is used. Reddy uses this last term in his PhD thesis on “Perceptually Modulated LOD”. (Reddy 1997)

In his thesis, Reddy summarizes the techniques used for controlling level of detail in VR, where 5 implementations of LOD are listed:

- Distance LOD
- Size LOD
- Eccentricity LOD
- Velocity LOD
- Fixed Frame Rate LOD

Area of Interest Management (AOIM) is a more containing word, which covers all means of LOD, including the Eccentricity LOD (i.e. foveation). In 2D image coding, this term corresponds to Region of Interest (ROI) and in eye-tracking systems Point of Regard (POR) has a similar hold.

In AOIM, by including the word “management”, the implication is moved from the area or the region, to the fact that something is done to manage it.

To summarize the umbrella of the terminology, the following bulleted list should also express relationship between these terms:

- Active Visualization

- Area of Interest Management (AOIM) (an active visualization method)
- Level Of Detail (LOD) (an AOIM method)

Terms used for human vision based LOD approaches:

- Gaze directed
- Gaze Contingent
- Perceptually Driven
- Eccentricity
- Foveated

Level of Detail management includes foveation but deals with other directions and ways the perception of detail varies as mentioned in the previous page. AOIM refers to the approaches that include LOD, while ROI and POR are used similarly as AOIM. All of these concepts are met when dealing with Active Visualization.

Note that, the term “perceptually-driven” in the literature covers a little more than foveal approach, it generally refers to a more generic account for what we visually percept, which, then includes distance-LOD and velocity-LOD and such. In many cases though, the term does not cover other senses – therefore it is not a complete reference to human perception.

“Active Visualization” is used similarly as Area of Interest Management (AOIM), although it implies a broader coverage. There are a number of other relevant and similar terms such as *smart sensing*, *visibility determination*, *attentional scene segmentation*, *active perception* – which are used to refer to the active visualization.

### 4. COMMON FOVEATION METHODS AND A FEW EXAMPLES

Can foveation be useful without eye tracking?

Before continuing with the methods and examples in the last chapter of this paper, it would be good to spare a few lines to explain why we are not concentrating on systems with eye-tracking equipment.

The foveation concept, by its description, is often closely related to tracking the eyes, or at least the head position of the viewer. The view-dependent real time rendering of the scene would require knowing the view point at all times, therefore it is obvious that one would want to have the eyes of the user tracked instantaneously.

But eye-tracking systems are complex and still very expensive. It is expressed that the expense is not because of the cost of the hardware, but because of the small market (UTEXAS 2002). Same source claims that when VR becomes more widely used, it would essentially include the tracking systems with it.

Until the cave VR experience is widely available to public, the notion of it can be passed via network-based interactive 3D environments, using the “worlds” that are created with VRML (Virtual Reality Modeling Language), Java3D and similar. There is currently no eye-tracking in experiencing the online interactive 3D models, nor they are fully immersive, but they certainly are available to a much larger audience. Along with their many other useful functions, they serve as a first introduction to VR and give an understanding of its potential.

In such systems, there is a bigger need for data management, compared to the offline systems. Therefore, it would be useful to test the foveation as a compression method for online VEs, for now, without assuming any eye tracking.

It is possible to track the head or eyes of the viewer also in front of an average computer monitor, if there is the necessary equipment, e.g. a camera following the user's head. This kind of systems based on a desktop computer, are known and called as "fishtank VR" in part of the VR literature. The needed accessories are not built in to the computers though, not just yet. When the purpose is to pass the notion and the experience of a VE to the public, obviously, minimum resources should be assumed.

When resources are at a minimum and we know that the eyes cannot be tracked, it is possible to work around the problem by tracking the pointing device. Cursor's position on the screen is known at all times, and one can often assume that the user is looking at where the pointer is. Particularly if the user is in an interactive environment, it is highly likely that she is interacting with an object by pointing at it, therefore she must be looking there. This knowledge is utilized in all 3D environments when a method such as visibility culling or Distance LOD is implemented.

For the Foveated LOD, or the Eccentricity LOD, instead of tracking the eye, interaction with the user by asking her to click at the point of interest, or chose an area of interest is used as the intermediate solution, like in the 2D image coding systems that will be mentioned later in this paper.

## 2D Foveation

Foveation has successfully been implemented as a space-variant image coding system by several groups or researchers. At this point, a brief look into a few examples and a review of some of the approaches in 2D applications will be given. There are a large number of publications that reports the geometry, the techniques and algorithms used for foveated image coding.

The common approach involves decoding the image with low-pass filtering methods into a foveation pyramid, and then decoding it back by up-sampling, interpolation and blending to make it into a smooth, displayable result image. (UTEXAS 2002, Chang 1998, Chang & Yap 1997, Kortum and Geisler 1996, Reddy 1997)

Use of Gaussian filters and wavelet compression techniques dominate the foveated space-variant image coding applications along with log-polar mapping. A C source code for a Gaussian implementation is published under General Public Licence (GPL) by Reddy (Reddy 1997) and a small application called "Foveator" solely for the image-foveation purpose can be found in (UTEXAS 2002). "Foveator" is compiled for Windows only and the source code is not published.



**Figure 3:** Some visual results using the "Foveator" program from UTEXAS. First image is original. The other two images are foveated, centered in where the arrows are pointing.

The notion of foveation is popular particularly for network applications when the images are large - as an alternative to typical progressive image coding algorithms. Naturally, 2D foveation algorithms are never concerned with depth, which is relevant for photogrammetry and VR.

## 3D Foveation: Eccentricity LOD

In most 3D applications, online or offline, it is acknowledged that a LOD management is useful and necessary. (Hoppe 1996, Reddy 1997, Luebke et.al. 2003) In general, there are three types of LOD; discrete, continuous and view-dependent (Luebke et.al.) and LOD algorithms consist of three major parts: generation, selection and switching. (Möller & Haines, 2003)

The applications that include foveation-like ideas are not too many in 3D, perhaps its obvious complexity makes it more computationally expensive than advantageous. But the concept is exhaustively studied in the literature. For instance, as mentioned earlier, Reddy introduces the "Perceptually-modulated LOD" for VR, therefore for 3D. A chapter in this work is aforementioned equal of foveation, the Eccentricity LOD. (Reddy 1997)

In another work, it is reported that "the system works with both 2D and 3D image datasets" - but in this case mentioned 3D datasets seem to be only images that does not have a vector model. (Schermann, 2001)

Basu et al. report a Java3D application using foveation: "The method combines foveating JPEG texture files with LOD representation in JAVA3D". In their work, although there is a vector wire frame on which the textures are mapped, the focus of foveation is on the texture mapping and not on the whole of the model. (Basu et.al. 2002)

In an earlier application, in 1993, Funkhouser had created a 3D architectural walkthrough model without photo-textures. He had re-created the model in several other resolutions, pre-computed the visibility information for each cell that he created under a spatial subdivision: all into a database system. It is a successful (though laborious) work in terms of "foveation" idea, which was called "visibility determination" in this work. (Funkhouser, 1993)

## 5. CONCLUSIONS & FUTURE WORK

Foveation is not utilized in photogrammetric visualizations as far as the author is aware. It certainly is interesting for all kinds of photogrammetric applications. Both in 2D and 3D visualizations, particularly on large screens (e.g. panoramic screens), when visualizing big aerial or satellite images, digital photogrammetric applications would benefit from such a reconstruction.

In a virtual environment, it would be ideal to have a photorealistic 3D world where foveation was possible and smoothly working based on a wireless eye-tracking system. VEs will eventually develop into fully perceptually modulated "places", also processing other human senses. Until then, testing the possibilities of foveation for web-based VR could make the interactive models accessible to an even wider audience.

The current methods such as mesh simplification, visibility culling, view frustum and view dependent rendering makes the high quality 3D graphics usable. Foveation is as useful as the listed methods.

No examples of stereo imaging systems were found utilizing foveation. The ongoing work for the author of this paper is to implement and test the performance of foveation for stereo imaging, in particular for anaglyph images, because viewing anaglyph does not require complex hardware acquisition and it can be published on the web.

## References

Chang, E., 1998. Foveation Techniques and Scheduling Issues in Thin wire Visualization. PhD Dissertation, Department of Computer Science, New York University.

Chang, E., Yap C., 1997. A Wavelet Approach to Foveating Images. Proc. 13th ACM Symposium on Computational Geometry, pp. 397-399, 1997.

Clark, J. H., 1976. Hierarchical geometric models for visible surface algorithms. Communications of the ACM 19(10), 547-554.

Basu, A., Cheng, I., Pan, Y., 2002. Foveated Online 3D Visualization. 16<sup>th</sup> International Conference on Pattern Recognition (ICPR'02), Canada, Volume 3, p.30944.

Funkhouser, T.A., 1993. Database and Display Algorithms for Interactive Visualization of Architectural Models. PhD thesis. Department of Computer Science, University of California at Berkeley.

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

Kalawsky, R.S., 1993. The Science of Virtual Reality and Virtual Environments. Addison-Wesley, Reading, MA.

Kortum, P., Geisler, W. 1996. Implementation of a foveated image coding system for image bandwidth reduction. SPIE Proceedings, 2657, 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/>.

Möller, T.A., Haines, E., 2003. Real-time Rendering. Textbook. A.K. Peters Ltd., 2nd edition, ISBN 1568811829 URL: <http://www.realtimerendering.com/>.

Murphy, H., Duchowski, A.T., 2001. Gaze-Contingent Level of Detail Rendering. EUROGRAPHICS 2001.

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. URL: <http://www.martinreddy.net/percept/>.

Schermann, 2001. 3D Foveated Visualization on the Web. PhD Thesis. Computer Science Department, University of Western Ontario.

## Web resources

Ware, C., Interactive Data Visualization Course Notes. URL: <http://www.ccom.unh.edu/vislab/VisCourse/VR.html>. Page visited on 26.11.2003.

UTEXAS 2002, Center for Perceptual Systems, The University of Texas at Austin URL: [http://fi.cvis.psy.utexas.edu/foveated\\_questions.htm#Q3](http://fi.cvis.psy.utexas.edu/foveated_questions.htm#Q3). Page visited on 26.11.2003.

## Acknowledgements

The author would like to acknowledge the financial support from Finnish Cultural Foundation.