STEREO FOVEATION

Arzu Çöltekin (Coltekin), Henrik Haggrén
Helsinki University of Technology, Institute of Photogrammetry and Remote Sensing

Arzu.Coltekin@tkk.fi, Henrik.Haggren@tkk.fi

ABSTRACT

This article presents an approach to area of interest management (AOIM) for stereoscopic images. The AOIM techniques are applied in different sub fields of computer graphics for 2D and 3D imagery and vector data. In this work, we take the concept and evaluate its potential for stereoscopic imaging, and create a simplistic implementation of the concept for test purposes to complete the evaluation.

Because of the multidisciplinary nature of the subject, we first present a background approaching the subject from different fields. After the background section, a description of our implementation, Foveaglyph, is given and some samples of stereo foveation is provided for a better understanding and analysis of the concept.

1. INTRODUCTION

The human visual system (HVS) is a complex biological system, which executes a number of highly demanding tasks in real time. The complexity of this system may be better appreciated if we quote the following facts: Visual intelligence occupies almost half of your brain's cortex (Hoffman, 2000), 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, 2004).

While the HVS is very complex, it is very efficient, and therefore the computer vision systems take inspiration from it. If cameras (e.g. robot eyes) are built in an effort to imitate what human eyes do to manage the visual input by e.g. moving the lenses; we consider this approach under the umbrella of *active vision*. It should be noted that active vision here refers not to sensing technology but to strategies for observation (Blake and Yullie (eds.), 1992), i.e. by altering camera's viewpoint.

HVS also inspires active visualization. This term is used (Chang and Yap, 1997, Blanke and Bajaj, 2002) to express that a conscious and planned effort was placed in planning how to visualize a scene. This scene may be a recorded image or video, or a 3D model, and any other graphic – what is essential is that it is managed visualized by digital equipment such as a computer, a display or a projector. A notable example of how HVS inspires active visualization is the use of a coordinate system that resembles the organization of human retina (Bernardino and Santos-Victor, 2002, Chang, 1998), referred to as log-polar mapping as it uses a polar coordinate system based on logarithmic graduation (Coltekin, 2006).

Among active visualization techniques, one is to determine the user's point or area of interest and render the level of detail in the scene accordingly. This is a level of detail management method based on the area of interest and it has been researched by computer vision and computer graphics communities. The method has been reported to be effective in a number of publications such as in Reddy, 1997, Perry and Geisler, 2002, Luebke et al., 2003, Linde 2003.

Based on the same principle, the image and video processing community has been working with a technique called foveation. As the name implies, foveation takes its inspiration from fovea, the small yellow spot on human retina. Fovea has a photoreceptor cell distribution that allows the eyes to receive the light in a controlled manner and creates different levels of detail in the perception of a scene. This in principle means sharp in the centre and less and less sharp towards the periphery.

Rods and cons are the photoreceptor cells that control different spatial and chromatic resolutions, i.e. cones capture high-resolution colour information while rods capture low-resolution monochromatic information. The following figure demonstrates the rod and con distribution in the eye, and in the fovea.

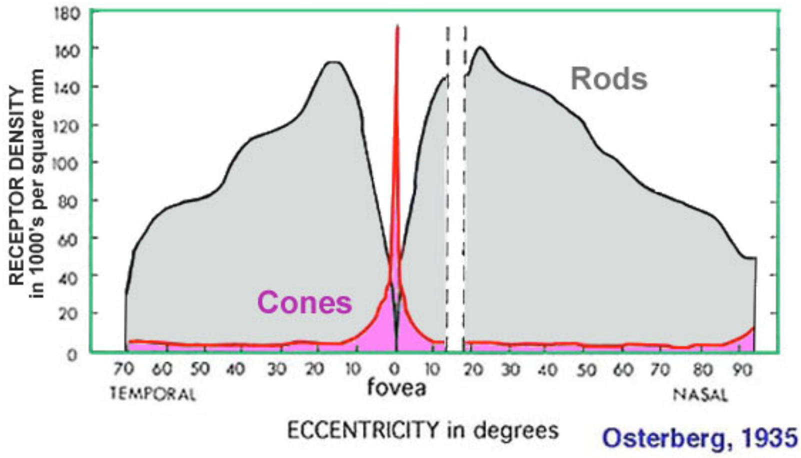


Figure 1: The fovea has a dramatically reverse rod and con distribution in the centre and in the periphery. Image reprinted from Osterberg (1935).

Foveation means, therefore, applying these principles in image management. It gives considerable compression and allows smart streaming when needed. Typically foveation is implemented for regular, two-dimensional images. This paper will give an overview of its potential for stereoscopic images.

Stereoscopic images provide 3D information about a scene. The most popular applications of stereoscopic imaging are for entertainment, amusement and educational purposes which are visualizations; these would include 3D cinema and in stereograms. In engineering, the stereoscopic images are often used for gathering 3D coordinates of a scene, and applications include but not limited to photogrammetry, robotics, telemedicine and virtual reality.

In most applications of stereoscopic imaging, real time processing and transmission of the data is required. Also when the images are static, i.e. there is no real time continuous transmission, 1:1 visualization of these images requires dealing with large datasets. These processes are demanding on the bandwidth and computational resources. Foveating the stereoscopic images will give us several advantages on managing the limited bandwidth and hardware resources as well as providing benefits with solving several problems associated with stereoscopic displays as indicated in a number of publications including Ware, 2004, Linde, 2004.

2. BACKGROUND

2.1 Area of Interest Management

Area of Interest Management (AOIM) depends on determining what the viewer is interested in at the time, when visualizing a scene. Once this is achieved, e.g. by means of eye tracking, or by means of an interactive interface, the scene, and its levels of detail would be rendered accordingly.

Level of detail management is not always based on determining the area of interest, techniques such as view frustum culling, occlusion culling are also LOD management and they are based on other technical factors, i.e. what is currently visible to the viewer. Size and distance based LOD management techniques are perceptually based, yet they are not based on the area of interest. In a geometric model, mesh simplification is applied considering what is visible to the user at the time, taking perceptual LOD factors into account, or based on area of interest.

The term LOD is used mainly by the computer graphics community, i.e. for geometric models, and the term foveation by the computer vision and image processing. Even though the conceptual link between the two is obvious, the literature on the two does not refer to one another extensively. We view foveation as a LOD management technique (Coltekin, 2006).

2.2 Foveation

Several approaches to foveation includes foveated lenses, dedicated hardware for foveating the images and software-based solutions. Software based foveation, according to Linde (Linde, 2004) may employ low-pass convolution masks variable in scale depending on the distance from the fixation coordinates, application of bank of filters with different cut-off frequencies (Lee et al., 1998) by the multiple application of a fixed size convolution mask, and more commonly, by the generation and subsequent pixel selection from a low-pass image pyramid (Perry and Geisler, 2002).

In this section we will refer to the examples of foveation software. Among several successful implementation reported in research papers, there is “Foveator” from University of Texas, a small, fast and effective foveation program that is available for download (see UTEXAS, 2006 for the URL for this program). The relevant publications are Kortum and Geisler, 1996; Geisler and Perry 1999; Perry and Geisler, 2002.

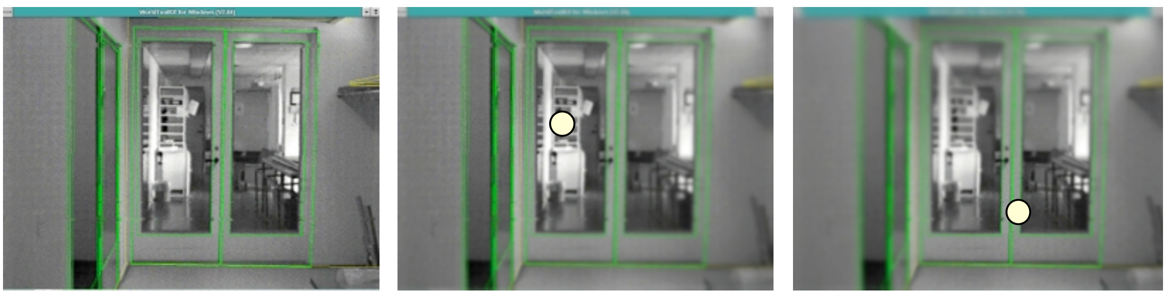


Figure 2: Visual results produced using the program called “Foveator” from University of Texas. The first image is the original. The other two images are foveated and approximate point of interest is marked with a circle.

Martin Reddy’s “Percept” is also a freely available, open source program that does 2D foveation (Reddy, 1997), though Reddy refers to 2D foveation as “Eccentricity LOD”, it describes exactly the same concept. Another significant work is from the New York University; a Java-based online foveation demo is available designed particularly geoinformation datasets in mind. The related work is published in Chang and Yap, 1997; Chang, 1998 and Chang et al., 2000.

2.3 Stereoscopic Perception and Stereoscopic Displays

Human depth perception is based on a number of “cues”; one of which is stereoscopic perception (also referred to as binocular vision). Stereoscopic perception allows us to distinguish fine details in depth up to a distance. This is a feature that typically the predators have. The ability to fuse two images into a single image is described in terms of the horopter, a circle in space defined by points that fall onto corresponding points on the two retinae. Points that lie on the horopter will be fused into a single image (Graham, 1951). Binocular vision can only occur where the fields of view of the two eyes overlap. The horizontal binocular visual field is about 120° out of a possible 200° (Pfautz, 2002).

Receiving two images, human brain fuses these images to create a fine 3D image. The fusion happens in a particular area called “Panum’s fusional area”. Named after Peter Ludvig Panum, a medical scientist who lived in 1820 to 1855. Panum’s fusional area is a model that defines the boundaries of binocular fusion. This is defined by finding the greatest horizontal disparity where fusion is still possible. When the disparity exceeds, i.e. when are outside the Panum’s fusional area, double vision (diplopia) will occur.

Several stereoscopic viewing techniques and stereoscopic displays have been developed based on the human stereoscopic vision over the years. The different stereoscopic viewing methods can initially be categorized by two major means; time parallel and time multiplexed methods. Time parallel methods send the two images at the same time, while time multiplexed (also referred to as time sequential or field sequential) methods send them in a sequence. A list of most common stereoscopic viewing and display techniques can be seen in Figure 3 (Coltekin, 2006).

While our example will use anaglyph method as it is one of the most platform-independent and least hardware dependent methods, autostereoscopic displays seem to be the most compelling type among others. This method has clear advantages as it does not require any additional equipment and could be viewed by several people, though it still is ideal only from a certain viewpoint. Recently a mobile phone and a laptop computer with optional autostereoscopic display was produced and marketed successfully by Sharp Inc. (Sharp 2005, Sanghoon et al., 2005).

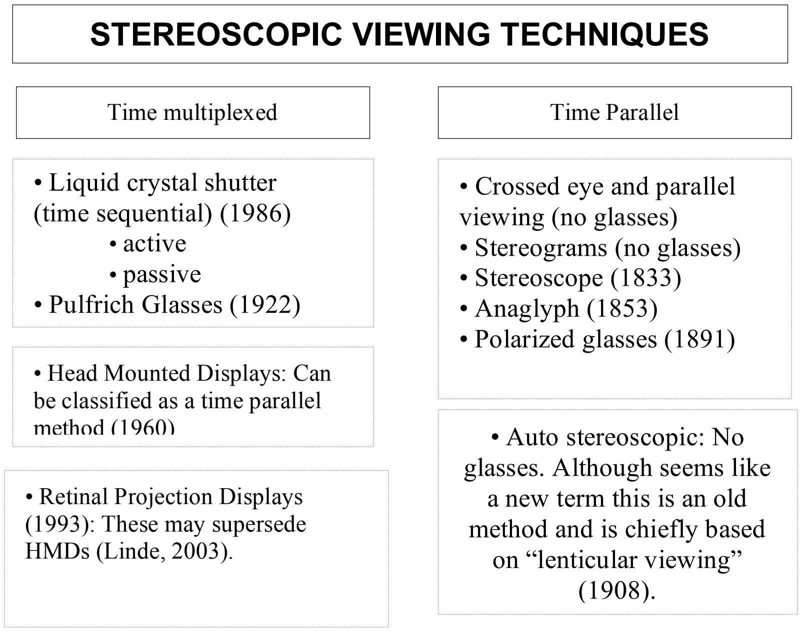


Figure 3: The methods of viewing stereoscopic images and graphics.
The years mentioned after the methods refer to their attributed invention times.

Based on our current understanding of stereoscopic perception and its latest mathematical models, there are successful binocular eye tracking systems, and gaze-contingent display (GCD) systems.

3. STEREO FOVEATION

Stereo foveation is typically for stereoscopic 1:1 viewing and most ideal for head mounted displays. If a real time implementation was to be done, it would be expected to perform identical to the HVS, in all domains including the spatial, chromatic and temporal. A real time foveated stereo visualization would heavily depend on the performance of the stereo eye tracker. Since we are currently analyzing the concept, we tested stereo foveation in an interactive environment, i.e. where user would chose the point of interest using a pointer.

Literature on stereo foveation is limited. There are several examples of depth of field (DOF) simulations (e.g. Mulder and van Liere, 2000, Krivanek et al., 2003, Piranda et al., 2005), but they are often done for monoscopic images.

Linde's focus/foveation (Linde, 2003 and Linde, 2004) combines a DOF simulation with a foveation procedure. He implements a DOF simulation on top of a 2D foveation utilizing *image z-buffer* for a synthetic VE, requiring a binocular eye tracker to determine the screen-incident coordinate for each eye. Linde works with JPEG images and utilizes an image histogram of the z-buffer to determine the spatial distribution of objects within the z-buffer.

Our implementation, *Foveaglyph*, takes a stereo image as input, and gives a foveated stereo anaglyph image as an output, based on the user specified point of interest. Anaglyph is the choice of stereoscopic viewing for its simplicity, as it requires minimum hardware.

The image matching is done using p2p (pixel-to-pixel stereo) by Stan Birchfield and Carlo Tomasi (Birchfield et al., 2005). All operations in a typical routine of stereo-foveation with *Foveaglyph* can be seen in Figure 4. The initial steps of camera calibration and lens distortion corrections are typical in photogrammetric tasks, however are not required by *Foveaglyph*.

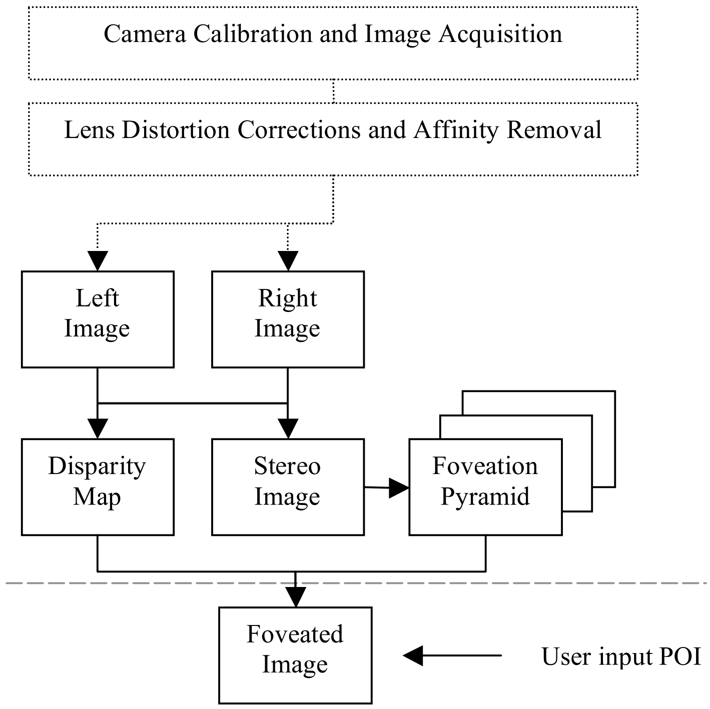


Figure 4: Foveaglyph operations (Coltekin, 2006).
The camera calibration and lens distortion corrections are optional for Foveaglyph.

The operations listed above the dotted line in Figure 4 are to be pre-computed whenever image pairs are available. This is chiefly because computation of the disparity map and creation of image pyramid may take a long time especially for large images. Computing them in advance is the typical practice and would allow real-time use of 3D foveation (Coltekin, 2006).

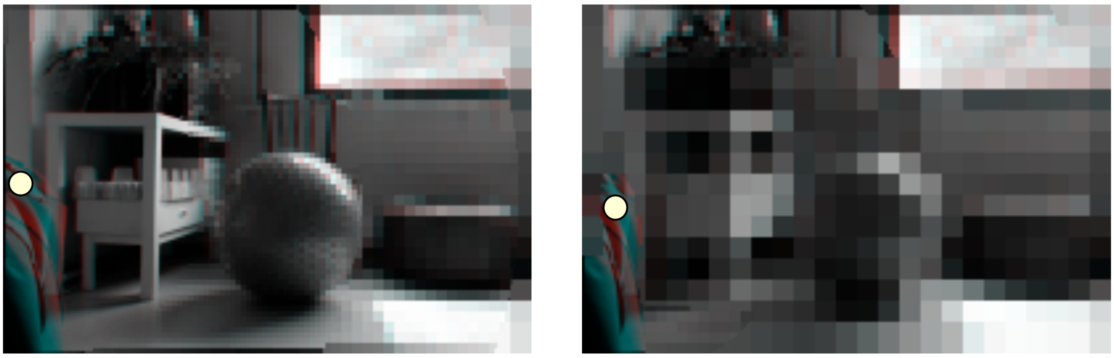
To study how the stereo foveation would work, we approximated a spherical geometric model and our foveation model is based on a step function. A spherical model and a step-function is a very simplified representation of the concept, however it fulfills its role, which is to act as a proof of concept.

3.1 Results

In the following figure (Figure 5), we present 2D and 3D results from *Foveaglyph*. While the visual results are exaggerated to demonstrate the effect, the numbers below the images tell us the compression that was provided with 2D and 3D foveations.

1) Original left image
200 x 132 pixels

2) Original right image
200 x 132 pixels



3) 2D fov., POI1 x=15 y=73, 4 levels, 72% gain

4) 3D fov., POI1 x=15 y=73, 4 levels, 98% gain

5) 2D fov., POI2 x=42 y=58, 3 levels, 62% gain

6) 3D fov., POI2 x=58, y=42, 4 levels, 70% gain

Figure 5: Results from Foveaglyph, for 2D and for 3D. The level switching is left visible in foveated images for demonstration purposes, it is possible to enhance this visual artefact easily with a smoothing filter.

These results are for the specific point of interests (POIs) and for the specified number of LODs. The position of the POI has an affect on the results. This is because when the image pyramid is created the most peripheral images have the least resolution, and when most of the foveated image elements come from the peripheral members of the image pyramid, the compression will naturally appear more aggressive.

4. CONCLUSIONS

We presented the potential of foveation for stereo imaging. Foveation, in 2D or in 3D is an effective compression method and it is “perceptually lossless” because the viewer’s point of interest is presented in full resolution, and the rest of the image obeys the same spatial degradation as the HVS. As the degradation goes unnoticed by the user in ideal case, we gain considerable space and time by applying foveation. By applying 3D foveation, compression gain is modest, depending on several factors, but consistent.

Our results indicate that with a carefully studied HVS-based model, stereo foveation may offer a number of benefits including computational ones listed above.

5. REFERENCES

Bernardino, A., Santos-Victor, J., 2002. A Binocular Stereo Algorithm for Log-polar foveated systems, 2nd workshop on Biologically Motivated Computer Vision, Tuebingen, Germany, Nov. 2002.

Blake, A., Yuille, A. (Eds.), 1992. Active Vision, ISBN 0-262-02351-2

Blanke, W., Bajaj, C., 2002. Active Visualization in a Multidisplay Immersive Environment Eighth Eurographics Workshop on Virtual Environments (2002)
S. Müller, W. Stürzlinger (Editors)

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

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

Chang, E.C., Mallat, S., Yap, C., 2000. Wavelet Foveation, Applied and Computational Harmonic Analysis 9, 312-335 (2000). Doi:10.1006/acha.2000.0324, available online at http://www.idealibrary.com

Coltekin (Çöltekin), A., 2005. A Visualization Method Based On Foveation. Proceedings of XXII International Cartographic Conference (ICC2005), A Coruna, Spain, 11-16 July 2005 ISBN: 0-958-46093-0

Coltekin (Çöltekin), A. 2006. Foveation for 3D Visualization and Stereo Imaging. Ph.D. thesis, Helsinki University of Technology, Helsinki, Finland. ISBN 951-22-8017-5, URL: http://lib.tkk.fi/Diss/2006/isbn9512280175/.

Geisler, W.S., J.S. Perry. 1999. Variable-resolution displays for visual communication and simulation. The Society for Information Display, 30:420-423.

Graham, C., 1951. Visual Perception. In S. Stevens (Ed.), Handbook of Experimental Psychology, pp. 868-920. New York: John Wiley & Sons.

Hoffman, D.D., 2000. 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

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.

Krivanek J., Zara J., Kadi B., 2003. Fast Depth of Field Rendering with Surface Splatting, In Proceedings of Computer Graphics International 2003. Los Alamitos: IEEE Computer Society Press, 2003, p. 196-201. ISBN 0-7695-1946-6.

Lee, S., Pattichis, M. S., Bovik, A. C., 1998. Foveated video quality assessment and compression gain, Technical Report UT-CVIS-TR-98-002, Center for Vision and Image Sciences, University of Texas at Austin.

Linde, van der, I., 2003. Space-variant Perceptual Image Compression for Gaze-contingent Stereoscopic Displays, PhD Thesis, Faculty of Science and Technology, Department of Computing, APU. Jan 2003

Linde, van der, I., 2004. Multi-resolution image compression using image foveation and simulated depth of field for stereoscopic displays, Proceedings of SPIE Volume: 5291, Stereoscopic Displays and Virtual Reality Systems XI Editor(s): Woods, Andrew J.; Merritt, John O.; Benton, Stephen A.; Bolas, Mark T. Published: 5/2004

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.

Mulder, J., van Liere, R., 2000. Fast PerceptionBased Depth of Field Rendering, Proceedings of the ACM symposium on Virtual reality software and technology, Seoul, Korea, Pages: 129 – 133, ISBN:1-58113-316-2

Ohshima, T., Yamamoto, H. and Tamura, H., 1996. Gaze-Directed Adaptive Rendering for Interacting with Virtual Space, Proceedings of the IEEE Virtual Reality Annual International Symposium (VRAIS), Santa Clara, CA, pp. 103–110.

Osterberg, G., 1935. Topography of the layer of rods and cones in the human retina. Acta Ophthal., suppl. 6, 1-103.

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

Pfautz, J.D., 2002. Depth Perception in Computer Graphics. Technical Report, University of Cambridge, Cambridge, UK, May 2000.

Piranda, B., de Sorbier, F., Arquès, D., 2005. Simulation of blur in stereoscopic image synthesis for virtual reality, WSCG 2005 posters. Proceedings of The 13th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision'2005, January 31-February 4, 2005, Plzen, Czech Republic. Copyright UNION Agency – Science Press

Reddy, M. 1997. Perceptually Modulated Level of Detail for Virtual Environments. Ph.D. Thesis. University of Edinburgh. URL: http://www.martinreddy.net/percept/

Sanghoon, L., Bovik, A.C., Kim, Y.Y., 2005. High Quality Low Delay Foveated Visual Communication over mobile channels. Journal of Visual Communication and Image Rrepresentation, Vol. 16 (2005) 180-211

Sharp 3D 2005. http://www.sharp3d.com/ last visited 30.7.2005

UTEXAS 2006. Center for Perceptual Systems, The University of Texas at Austin
URL: http://svi.cps.utexas.edu/software.htm, last visited on 21.08.2006.

Ware, C. 2004. Information Visualization - Perception for Design. Second edition. Elsevier Inc., Morgan Kaufmann Publishers. ISBN 1-55860-819-2