Reevaluating Stereo and Motion Cues for Visualizing Graphs in Three Dimensions

Colin Ware* and Peter Mitchell#
Data Visualization Research Lab. CCOM, University of New Hampshire

Abstract
It has been known for some time that larger graphs can be interpreted if viewed in 3D than in 2D. Both kinetic depth cues and stereoscopic depth cues increase the size of the structure that can be interpreted. However, prior studies were carried out using displays that provided a level of detail far short of what the human visual system is capable of resolving. This is especially problematic because human stereoscopic vision is known to be a super-acuity, it operates best under conditions where fine details are present. Therefore we undertook a graph comprehension study using a very high resolution stereoscopic display. We examined the effect of stereo, kinetic depth and using 3D tubes versus lines to display the links. The results showed a much greater benefit for 3D viewing than previous studies. For example, with both motion and depth cues, unskilled observers could see paths between nodes in 333 node graphs with a better than 10% error rate. Skilled observers could see up to a 1000 node graph with less than a 10% error rate. This represented an order of magnitude increase over 2D display. These findings are discussed in terms of their implications for information display.

CR Categories: 1.3.6 [Computer Graphics]: Methodology and Techniques. H.5.2 [Information Interfaces and Presentation]: User Interfaces-evaluation/methodology

Keywords: graph drawing, information visualization, 3D visualization, node-link diagrams.

1 Introduction

One of the basic tenets of information visualization is that if information structures can be visualized then they may be interpreted more easily. Unfortunately, there are limits to the size and complexity of structures that can be displayed on a 2D display; for example, most node-link diagrams that are produced for various branches of computer science have fewer than 30 nodes and a similar number of links between them. Although some very large node-link diagrams have been shown [e.g. Munzner, 1997] the goal has been to give an impression of the overall structure, rather than allow people to see individual links. To be sure, a number of interactive techniques allow users to rapidly browse information networks that are much larger than can be placed on a single computer monitor. For example the cone tree [Robertson and Mackinlay, 1993] showed large trees in 3D but required users to rotate various levels of the tree to find the node they were seeking. Other techniques allow users to interactively highlight or extract subgraphs of a larger graph and thereby provide interactive access to the whole [Munzner et al, 1999; Wills, 1999; Ware and Bobrow, 2005]. However, if information can be perceived without any manipulation of the display this should generally make for a more rapid understanding because interaction via a computer mouse will always take more time than making an eye movement. Thus the question of how large a graph can be seen in a non-interactive display is an important one; interactive techniques may always be added to increase the usable graph size still further.

*email:cware@ccom.unh.edu
#email: pete.mitchell@comcast.net

It is well known that larger network structures can be seen in 3D, where "in 3D" means that stereoscopic viewing and or kinetic depth cues are provided [Sollenberger and Milgram, 1993; Ware and Franck, 1996]. However, it is also the case that studies investigating the value of 3D displays have been done with conventional monitors having display resolutions considerable less than the eye can see. There are reasons to think that having high resolution is particularly important for looking at 3D structures and for this reason we decided to revisit the question of how much can be seen with and without 3D viewing, and therefore we carried out this study with a very high resolution stereoscopic display.

The task we chose to investigate is that of tracing paths in graphs, and our question is "How large a graph can we display and still see paths linking nodes?"

2 Perceptual Issues

Perception researchers consider the problem of perceiving distance from the viewpoint in terms of *depth cues*. The following is a list of some of the more important ones:

1) Stereoscopic disparities	5) Occlusion
2) Kinetic depth	6) Shape from shading.
3) Perspective	7) Others – cast shadows, focus, eye convergence
4) Texture and size gradients.	

Stereoscopic disparity and kinetic depth are likely to be the most important depth cues for looking at 3D node-link diagrams. To see why the others are less relevant we briefly review them. First, perspective projection is a good cue if there are parallel lines in a 3D scene; these converge to a "vanishing point" ; but in a 3D graph the lines will be arbitrarily oriented, and so perspective will provide little or no information. Supporting this, Ware and Franck [1996] found no significant difference between a graph viewed in perspective and one with an orthographic projection for a path tracing task.

Unless the links of a graph are rendered as solid tubes, shading and occlusion will provide little information, and unless they have textured surfaces, texture gradient information will not be available. If tubes rather than lines are used, they must necessarily be thin to allow a large number of links to be made clear, but when the tubes are thin they are unlikely to convey much shading information.

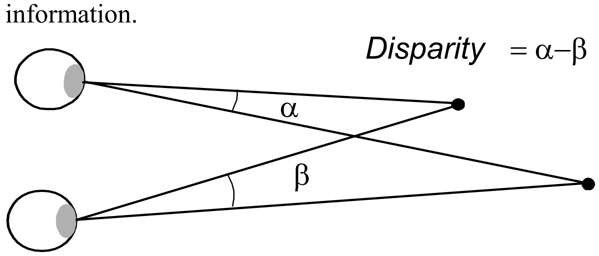


Figure 1. Stereopsis is based on the angular difference between pairs of points in the visual field. These differences are called disparities

A graph can be rendered to show cast shadows on the ground plane and this has been used in the case of cone trees to make the structure clearer [Robertson et al, 1993]. However, for shadows to provide useful information a perceptual correspondence must be established between the shadow and the link or node casting the shadow. With larger graphs this is likely to be impossible. Finally, focus and eye convergence are weak depth cues in general [Howard and Rogers, 1995] and there is no reason to think that they would help people perceptually trace paths in 3D graphs.

Thus we are left with stereoscopic disparities and kinetic depth. These are the cues that are most likely to be useful for perceiving 3D graphs and trees and they have received the most attention from researchers.

2.1 Stereoscopic viewing

Stereoscopic depth relies on the detection of relative differences, called disparities, between pairs of features imaged in the two eyes. Figure 1 illustrates.

Our ability to see stereoscopic depth allows for extra-ordinarily fine judgments. For example, Tyler [1975] found that acuity for discriminating a wavy line varying in depth was better than 1 arc second (best for a wave period of 1 cycle per degree). This is much better than could be predicted from the size of retinal receptors and indicates that the visual system must integrate the signals from multiple receptors. However other patterns have other limits. The threshold for detecting line disparity is about 12 arc seconds [Howard and Rogers, 1995]. However, stereoscopic orientation of featureless lines is not well specified by a stereo pair [van Ee and Schor, 2000] and in the case of a node link diagram the most important depth information may be provided by the nodes. The extreme sensitivity of the human visual system to disparities is the reason for using the very high resolution display chosen for this study.

2.2 Structure-from-motion cues

The projected image of a rotating 3D wire object appears strongly three dimensional, even though when the motion is stopped the object appears completely two dimensional. This is called the kinetic depth effect [Wallach and O'Connell, 1953]. This is one of several structure-from-motion depth cues and it relies on a built-in assumption by the visual system that objects are rigid.

Studies have compared the relative value of stereoscopic depth and motion parallax for a variety of tasks. The results make it clear that when considering the value of different depth cues it is essential to take the precise task into account. Consider the following examples: for the task of surface shape perception, stereo and motion cues appear to be roughly equivalent [Norman et al, 1996] although this may depend on the shape of the objects being observed and for how long. For cylindrical objects under stereoscopic viewing it is easier to resolve curvature differences for horizontal cylinders than for vertical cylinders [Rogers and Gagnello, 1989]. Concerning the viewing time, a study by Uumori and Nishida [1994] showed that for random dot surfaces, motion parallax was initially the dominant cue but after a few seconds stereoscopic depth became dominant. A study of the perception of the orientation of real twig objects [Frisby et al,1966] found that stereoscopic depth cue was more important than motion. This may have been at least partially due to the presence of fine visual textures on the surfaces of the twigs.

The particular task we are interested in is tracing paths in graphs. Studies of both tree

[Sollenbenberger and Milgram, 1993; Arthur, et al, 1993], and graph structures [Ware and Franck 1996] have found that motion is a more important cue than stereopsis. Ware and Franck [1996] found roughly linear increases in errors with graph size, but with different gradients for different viewing conditions. Their task was to determine the presence or absence of a path of length two between two highlighted nodes. The results showed that adding stereoscopic depth allowed for a graph 60% larger to be perceived, adding motion parallax allowed for a graph 120% larger to be perceived and adding both allowed for a graph 200% larger to be perceived. To give a specific example, they found that for an error rate of 20% approximately 55 nodes could be seen in 2D, but when viewed in 3D with stereo and motion parallax information a graph of 160 nodes could be viewed.

However, all of these studies used conventional monitors (1024x768 resolution) and frame-sequential shutter glasses as the display device. Because stereoscopic depth is a super acuity it is possible that they may have considerably underestimated the importance of stereopsis in perceiving large structures and they may have considerable underestimated the size of the largest structure that can be clearly viewed.

A criticism that has been leveled against the previous [Ware and Frank, 1996] study was that the layout of the graph was random. They justified this by arguing that random layout, favors neither 2D or 3D viewing. Nevertheless, in practice random layout is not used in graph visualization.

The present study was designed to address both of the display resolution issue and the layout issue. We used a display capable of displaying images at the limit of the resolution of the human eye. It had 9.2 million pixels for each eye and we also anti-aliased critical parts of the display. Thus we can claim to be addressing the question of how large a graph can be seen in a way that is not constrained by spatial resolution (although it may have been constrained by temporal resolution). We chose to use spring layout graphs since spring layout is widely used in practice [di Battista, 1999].

The spring layout algorithm involves representing graph edges as springs so that connected nodes are pulled together if they are further apart than the resting length of the spring and pushed apart if they are closer. At the same time all nodes repel one another according a function that is proportional to the inverse of the distance between them. An iterative process is applied until the system reaches equilibrium. Spring layout can be readily done in either 2D or 3D. Thus we are able to compare 3D spring layout with and without stereo and motion cues with 2D layout. In addition we also decided to compare graphs rendered with the edges represented by 3D tubes with graphs that used lines for edges.

3 Method

Our method was to have users visually trace paths in spring layout graphs to determine if there was a path or length 2 or 3 between two highlighted nodes. Our dependent measures were error rate and time to respond.

3.1 The Display

Our display is illustrated in Figure 1. It consists of a Wheatstone mirror stereoscope [Wheatstone, 1938]. This use of front surface mirrors has the advantage that there is no ghosting or reduction in image brightness, a problems that plagues many stereoscopic display technologies. The displays were Viewsonic VP 2290b monitors. Each of these displays has 3840x2400 pixels. With a display area 47.7 x 29.7 cm giving an individual pixel size of 0.0125 cm. The screens were set at a viewing distance of 105 cm. This yielded a visual angle per pixel of approx 24 seconds of arc. This is comparable to the size of receptors in the fovea and is easily sufficient to display the finest grating pattern that can be resolved by the human eye – about 60 cycles per degree [Campbell and Green, 1965]. The displays were driven by four PCs each containing an NVidia Quadro FX 3000G card and an AMD Athalon FX51 3400 processor. Each card supplied images to half of each display.

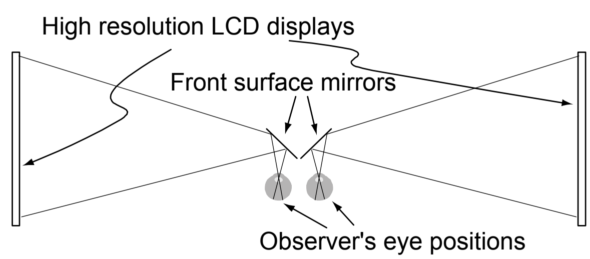


Figure 2. This diagram illustrates the Wheatstone stereoscope arrangement.

Figure 3. A subject viewing the display.

3.2 The Task

On each trial the subject was given a graph with two of the nodes highlighted in red. The subject's task was always to determine if the nodes were linked by a path of length 2 or 3. The subject pressed the left mouse button if the answer was 2 and the right mouse button if the answer was 3. Each viewing condition was displayed for a maximum of 5 seconds, after which the screen went blank until the participant responded. The reason for the time limit was that we were interested in visual searches that could be conducted rapidly, as opposed to those requiring laborious visual searches.

3.3 Conditions

There were four conditions with 3D spring layout of the graph and one with 2D spring layout.

- **No stereo, no motion.** Participants saw a static, non stereo perspective projection of the 3D graph **Stereo.** Participants viewed the graphs as a stereo pair.
- **Motion.** Participants viewed the graph rotating smoothly at a rate of one complete cycle every 36 seconds.
- **Stereo and Motion.** Participants viewed a graph with both stereo and motion cues.
- **2D layout.** The graph was layed out in 2D. This means that there was no occlusion of one node by another.

There were two rendering styles.

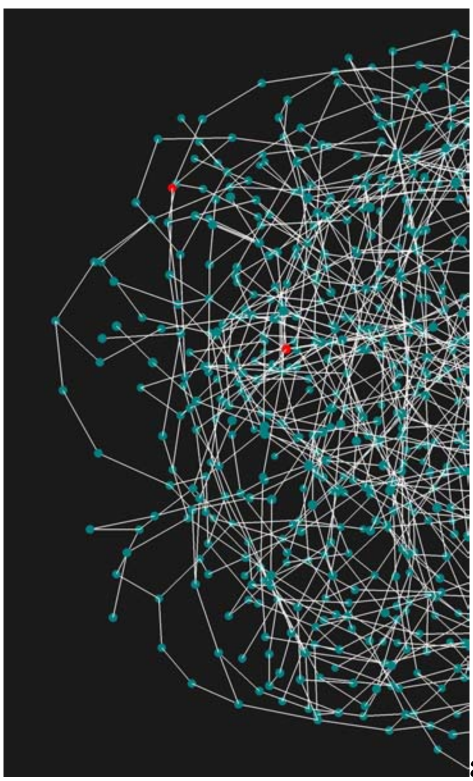
- **Lines.**
- **3D tubes.**

There were four different graph sizes: **33, 100, 333** and **1000** nodes.

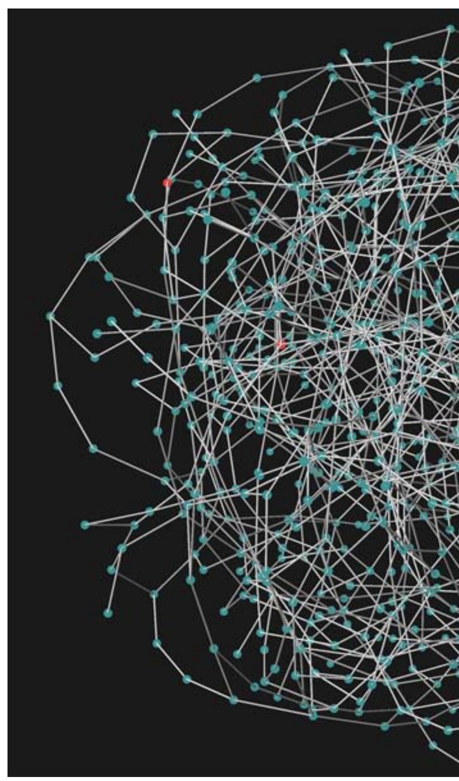
3.4 The Graphs

The algorithm randomly assigned links in such a way that the following statistics resulted.

6% of the nodes had degree one (leaf nodes); 37% had degree two ; 45% had degree three; 10% had degree four; <2.5% had degree five.



a



b

Figure 4. Half of a 1000 node graph drawn (a) with lines for the links and (b) with solid tubes for the links.

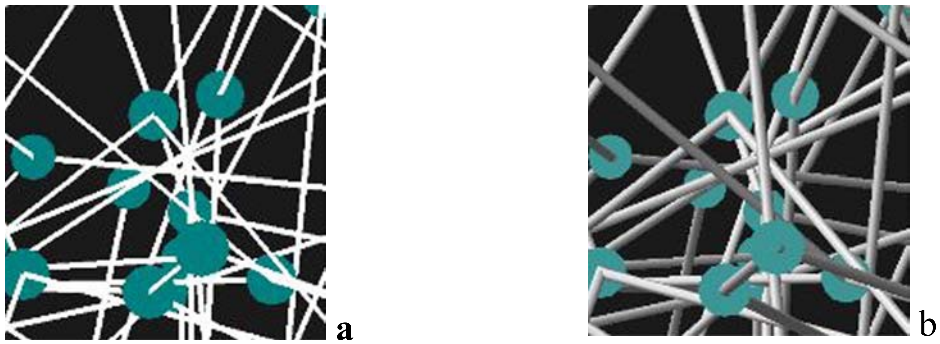


Figure 5. Snippets of the 1000 node graph roughly 2x actual size. (a) line rendering. (b) tube rendering. Should be viewed from 2 m to get equivalent visual effect.

The graphs were laid out using a spring forces iteratively applied [di Battista et al 1999]. There were three kinds of forces used in the layout: nodes repelled each other with a force inversely proportional to the square of the distance between them. Nodes connected by an edge were subjected to a force proportional to the deviation from an edge separation constant (2.4 cm in 3D, 1.2 cm in 2D). In order to make a more compact overall structure, nodes were subjected to independent forces along the x, y and z axes towards the origin and proportional to the cube of the distance from the origin.

The most computationally costly aspect of graph layout is the $O(n^2)$ cost of computing the repulsion forces between nodes. To accelerate the layout we used a 3D grid extending the method reported by Fructerman et al [1991]. By overlaying a 5x5x5 grid over the display volume we could easily keep track of which nodes fell into which grid cells (cubes) as well as how many nodes were in each cell. When calculating repulsion forces for a given node, any node in the same grid cell or an immediately adjacent grid cell was applied directly. For every other (non-neighboring) grid cell, a single force was applied from the center of the cell, weighted by the number of nodes it contained. This resulted in a comparable application of repulsion forces while substantially reducing the time required to stabilize a spring layout for the larger graph sizes.

One problem that we encountered in the pilot phase was that the nodes separated by a path of length 3 were on average, more widely separated than nodes separated by a path of length 2. This provided a cue for the response that had nothing to do with the viewing condition. In order to remove this extraneous variable we selectively selected paths in such a way that the mean Euclidean distance between start nodes and end nodes was the same.

All graphs were rendered against a black background. Line edges were drawn on a line thickness of 3 pixels. Cylinder edges had a radius of 0.036 cm. The node diameter was 0.3 cm. In the Motion conditions the entire graph rotated about the y-axis at a rate of 0.5° per frame @ 20 fps = 10°/sec. Highlighted nodes were rendered red (rgb(1,0,0)) while all other nodes were rendered cyan (rgb(0,0.5,0.5)). Examples of 1000 node graphs are given in Figures 4 and 5.

3.5 Participants:

The participants were 15 undergraduate students, paid for participating. In addition, the two authors of this paper also carried out the experiment to get an estimate of performance from more experienced observers.

3.6 Procedure

The experiment was the product of the 5 viewing conditions with 2 rendering styles and 4 graph sizes, yielding 40 different conditions. Trials were given in blocks of 20 for each condition. The entire set of conditions was randomly ordered. At the start of the experiment participants were given a training session where they were given a few trials in each of the 5 viewing conditions with both small sized and large sized graphs.

4 Results

We dropped one of the participants from the analysis because of error rates exceeding 30% even in the easiest (small graph) condition. The analysis was carried out on the remaining 14 participants. Figure 6 summarizes the main error rate results. As can be seen, the combined stereo and motion condition yielded the lowest error rate. Having either stereo or motion was the next best. There was little difference between 2D layout and 3D layout (viewed without 3D depth cues) except in the middle range of 100-333 nodes. An ANOVA on condition, number of nodes, and rendering style revealed the following effects.

There was a main effect for number of nodes (F(3,39) = 21.6; p< .001) and a main effect for condition (F(4,52) = 43.2; p < 0.001) There was also significant effect for whether tube rendering or line rendering of links was used. F(1,13)= 15.03; p < 0.02). There were about 2.5% more errors overall with tubes than with lines.

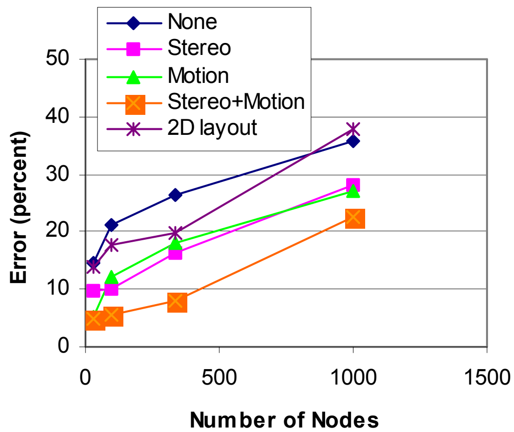


Figure 7. Errors as a function of graph size. Averaged data from 14 participants.

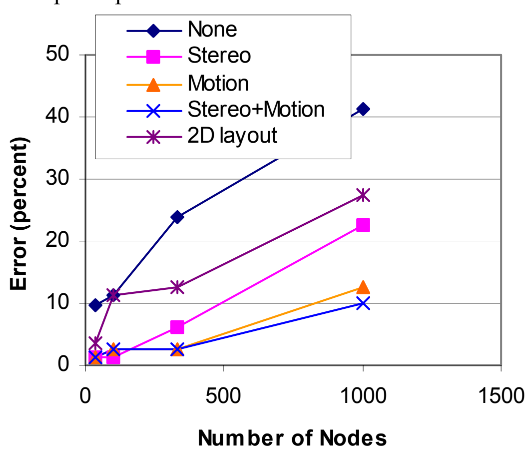


Figure 8. Errors as a function of graph size. Averaged data from 2 experienced participants.

The error rate results from the two authors results are summarized in Figure 7. As can be seen we had considerably lower error rates than the inexperienced participants. In the stereo+motion condition we achieved 10% or better errors even with the 1000 node graph.

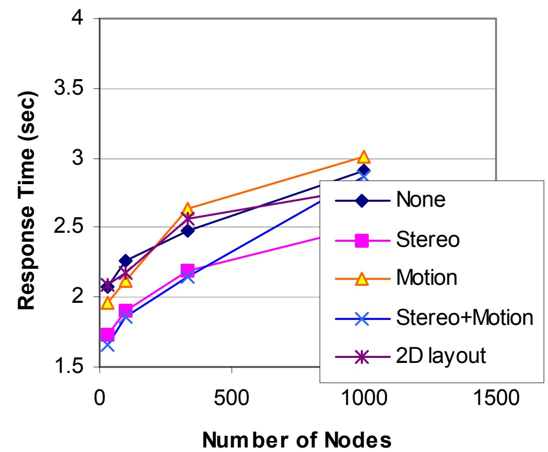


Figure 9. Time to respond as a function of graph size for the different conditions. Averaged data from 14 inexperienced participants.

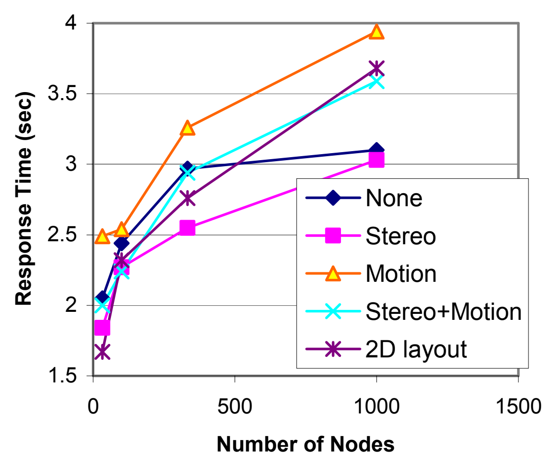


Figure 10. Time to respond as a function of graph size. Averaged data from 2 experienced participants.

Figure 9 summarizes the time to respond data as a function of graph size for the 14 inexperienced participants. Unsurprisingly this shows that response times increase as a function of graph size. It is also apparent that the stereo conditions resulted in the shortest response times. An analysis of variance was carried out on the factors of graph size vs viewing condition (stereo, motion, etc) . This revealed a main effect for graph size (F(3,39) = 36.9; p < 0.001) and a main effect for condition (F(4,52) = 3.85; p < 0.01) There was no interaction between them. A post hoc Tukey test for honestly significant difference applied to viewing condition revealed two groups: one group containing the stereo and the stereo+motion conditions and the other contained the non-stereo viewing conditions. Average response times were 15% faster with stereo than without (2.1 sec vs 2.42 sec).

Figure 10 shows the time to respond data for the experienced observers. Overall our responses took longer than the inexperience observers especially with the larger graphs. Response times were also the quickest for the stereo condition, however the stereo+motion was not as fast.

5 Discussion

From a practical point of view, the most striking aspect of our results is that 3D depth cues allowed participants to see paths in graphs containing 333 nodes with better than 92% accuracy. More experienced observers were able to see graphs up to 1000 nodes with better than 90% accuracy. With 2D viewing and 2D spring layout a 33 node graphs yielded comparable error rates. Thus, we find roughly an order of magnitude increase in the size of the graph that can be "read" (where we consider "reading" to be the identification of short paths) when 3D viewing is available using stereo and motion depth cues. These gains are dramatically better than those reported previously by Ware and Franck [1996] who only reported a 3-fold gain from 3D viewing. The author's data differed from that of the inexperienced observers in that for us motion was the most useful cue, not stereo. However, stereo viewing produced more rapid responses for both groups of observers.

We attribute the difference between the inexperienced and the experienced participants mostly to differences in motivation. The experiment was quite long and monotonous and producing 640 careful responses requires a considerable commitment. The reason why the experienced observers produced lower error rates was probably due partly to the fact that we took longer. This may also account for the fact that the experienced observers appeared to benefit more from motion cues. Motion produces a wider range of views than stereoscopic viewing if sufficient time is taken to wait for a particular pathway to be revealed.

There are a number of factors that differentiate this study from the previous work. First the resolution of the display was much higher. Second, the nature of the graphs themselves was very different. In Ware and Frank [1996] the graphs were randomly laid out and consisted of a large number of small connected components. These graphs were fully connected and given spring layout. Third, the update rates differed, however, the present experiment actually had lower update rates (20 Hz) than the previous one (30 Hz). Fourth, is the difference between frame sequential stereoscopic display and the use of a mirror-based Wheatsone stereoscope. Frame sequential displays usually have some ghosting, particularly for bright lines on a dark background. A mirror stereoscope has no ghosting at all. We are not able from this single study to differentiate between these factors although we are currently engaged in an experiment where we systematically vary the resolution. We are also working on improving the layout to attempt to increase the size of the graph that can be clearly viewed by both inexperienced and experienced observers.

We wish to add a phenomenological observation to our results. The high resolution display was much easier to view than the low resolution display. We can attest that both spatial and temporal aliasing made the lower resolution displays unpleasant to use for extended period. In contrast the high resolution display seems no more visually stressful to view than the real world although the drawback of a mirror stereoscope is that the head must be held in a constant position.

Concerning the practical value of 3D display of graphs, the sheer magnitude of our results suggests that there may be value is displaying social nets or communication nets in this way. We make no claims regarding data is not representable as node-link diagrams but node-link diagrams are very widely used in practice therefore we have hopes that 3D display may become as useful tools in data analysis.

Acknowledgements

We are grateful to Roland Arsenault who wrote the software that supported stereoscopic viewing on the five computer system. He also wrote the code that did anti-aliasing. Funding for this project was provided by NSF ITR grant 0324899 to the first author.

References

ARTHUR, K, BOOTH, K.S. AND WARE, C. 1993. Evaluating Human Performance for Fishtank Virtual Reality. *ACM Transactions on Information Systems,* 11(3), 239-265.

CAMPBELL, F.W. AND GREEN, D.G. 1965. Monocular versus binocular visual acuity. *Nature,* 208, 191-192.

FRISBY, J.P., BUCKLEY, D., AND DUKE, P.A. 1966. Evidence for recovery of lengths of real objects seen with natural stereo viewing. *Perception,* 25, 129-154.

DI BATTISTA G, EADES P, TAMASSIA R, TOLLIS IG. *Graph Drawing: Algorithms for the Visualization of Graphs*. Prentice-Hall Englewood 1999; Cliffs, NJ.

FRUCTERMAN T, REINGOLD E. 1991. Graph drawing by force-directed placement, *Software Practice and Experience* 21(11): 1129-1164.

HOWARD, I.P. AND ROGERS, B.J. 1995. *Binocular Vision and Stereopsis.* Oxford Psychology Series No 29. Oxford University Press. Oxford.

HUANG ML, EADES, P, WANG, 1998. Online animated visualization of huge graphs using a modified spring algorithm. *Journal of Visual Languages and Computing* 9(6): 623-645.

KLEIBERG E, VAN D WETERING VAN WIJK, JJ. 2001. Botanical visualization of huge hierarchies. *IEEE Symposium on Information Visualization,* (San Diego, CA), 87-94.

VAN EE, R., AND SCHOR, C.M. 2000. Unconstrained stereoscopic matching of lines. Vision Research, 40, 151-162.

MUNZNER, T., GUIMBRETIÈRE, F AND ROBERTSON, G. 1999. Constellation: A visualization tool for linguistic queries for MindNet. Proc. IEEE Symposium on Information Visualization. October 25-26 1999, San Francisco, CA, 1999, 132-135.

MUNZNER, T. 1997. H3: Laying out Large Directed Graphs in 3D Hyperbolic Space. Proc. IEEE Symposium on Information Visualization. 2-10.

NORMAN, J.F. TODD, J.T. PEROTTI, V.I AND TITTLE, J.S. 1996. The visual perception of 3-D length. Journal of Experimental Psychology: Human Perception and Performance. 22, 173-186.

PARKER, G., FRANCK, G., AND WARE, C. 1998. Visualization of Large Nested Graphs in 3D: Navigation and Interaction. *Journal of Visual Languages and Computing.* 9, 299-317.

PLAISANT C, GROSJEAN J, BEDERSON BB. 2002. SpaceTree: Supporting exploration in a large node-link tree, design evolution and empirical evaluation. *IEEE Symposium on Information Visualization* (Boston, MA). 57-64.

ROBERTSON, G., MACKINLAY, J.D. AND CARD, S.W. 1993. Information Visualization using 3D interactive animation. *Communications of the ACM.* 36(4) 57-71.

ROGERS, B. AND CAGNELLO, R. 1989. Disparity curvatures and the perception of three-dimensional surfaces. *Nature* 339: 137-139.

SOLLENBERGER, R.L. AND MILGRAM, P. 1993. The effects of stereoscopic and rotational displays in a three-dimensional path-tracing task. *Human Factors* 35(3) 483-500.

TYLER, C. 1975. Spatial organization of binocular disparity sensitivity. *Vision Research* 15, 583-590.

UUMORI, K. AND NISHIDA, S. 1994. The dynamics of the visual system in combining conflicting KDE and binocular stereopsis cues. *Perception and Psychophysics*, 55(5), 526-536.

WALLACH H., AND O'CONNELL, D.N. 1953. The kinetic depth effect. *Journal of Experimental Psychology*, 45: 205-217.

WARE, C. ARTHUR, K. AND BOOTH, K.S. 1993. Fishtank Virtual Reality. *INTERCHI'93 Proceedings* 37-42.

WARE, C. AND BOBROW, R. 2005. Supporting visual queries on medium-sized node-link diagram. *Information Visualization.* 4, 49-58.

WARE, C. AND FRANCK, G. 1996. Evaluating Stereo and Motion Cues for Visualizing Information Nets in Three Dimensions. *ACM Transactions on Graphics*.15(2) 121-139.

WILLS GJ. 1999. NichWorks: Interactive visualization of very large graphs. *Journal of Computational and Graphical Statistics* 8(2) 190-212.

WHEATSTONE, C. 1838. Contributions to the physiology of vision. Part the first. On some remarkable and hitherto unobserved phenomena of binocular vision. *Philosophical Transactions of the Royal Society.* 128, 371-394.

WILLS, G.J. (1999) NicheWorks: Interactive Visualization of very large graphs. *Journal of computational and Graphical Statistics,*8(2) 190-212