Visualizing Graphs in Three Dimensions

COLIN WARE and PETER MITCHELL
University of New Hampshire

It has been known for some time that larger graphs can be interpreted if laid out in 3D and displayed with stereo and/or motion depth cues to support spatial perception. 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. Therefore, we undertook a graph comprehension study using a very high resolution stereoscopic display. In our first experiment, we examined the effect of stereoscopic display, 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 stereoscopic depth cues, unskilled observers could see paths between nodes in 333 node graphs with less than a 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. In our second experiment, we varied both nodes and links to understand the constraints on the number of links and the size of graph that can be reliably traced. We found the difference between number of links and number of nodes to best account for error rates and suggest that this is evidence for a "perceptual phase transition." These findings are discussed in terms of their implications for information display.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Evaluation/ methodology, Theory and method*

General Terms: Design, Experimentation, Human Factors

Additional Key Words and Phrases: Visualization, network visualization, graph visualization, stereoscopic displays

ACM Reference Format:

Ware, C. and Mitchell, P. 2008. Visualizing graphs in three dimensions. ACM Trans. Appl. Percpt. 5, 1, Article 2 (January 2008), 15 pages. DOI = 10.1145/1279640.1279642 http://doi.acm.org/10.1145/1279640.1279642

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 two-dimensional (2D) display; for example, most node–link diagrams 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. One way of increasing the size of the graph that can be understood is

Funding for the first experiment was provided by NSF ITR grant 0324899 and for the second experiment was provided by the Advanced Research and Development Activity (ARDA) under contract number HM1582-05-C-0023.
Authors' address: Colin Ware and Peter Mitchell, Data Visualization Research Lab, Center for Coastal and Ocean Mapping, University of New Hampshire, Durham, New Hampshire 03824; email: cware@ccom.unh.edu

through interactive techniques, allowing 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 et al. 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; Plaisant et al. 2002; Ware and Bobrow 2005]. However, if information can be perceived without any interaction, this will 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 noninteractive display is an important one; interactive techniques may always be added to increase the usable graph size still further.

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 considerably 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. Therefore, we carried out this study with a very high-resolution stereoscopic display.

There are many cognitive tasks that are supported viewing node–link diagrams and these may involve understanding either its large-scale or small-scale structure. In our uses of graphs, we are concerned with the case when a data analyst wishes to know the near neighborhood of a particular node. For example, in reasoning with a social network diagram, people are usually concerned with *who knows who* (one link path) or occasionally with near acquaintance (two link paths). Similarly in using a software diagram it is usually the case that a programmer is concerned with one or two link paths (e.g., which procedures, and variables, does this entity use directly). Considerations, such as perceiving the symmetry of the graph, are less important in these kinds of applications. Accordingly, 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
2. Kinetic depth
3. Perspective
4. Texture and size gradients
5. Occlusion
6. Shape from shading
7. Others—cast shadows, focus, eye convergence

Stereoscopic disparity and kinetic depth are likely to be the most important depth cues for looking at three-dimensional (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. Second, unless the links of a graph are rendered as solid tubes, shading and occlusion will provide

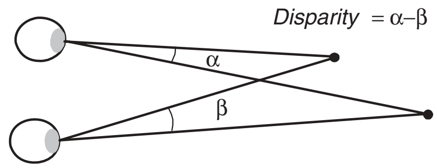


Fig. 1. Stereopsis is based on the angular difference between pairs of points in the visual field. These differences are called disparities.

little information. 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 useful shading information. Third, unless links have textured surfaces, texture gradient information will not be available.

Fourth, a graph can be rendered to show cast shadows on the ground plane; 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; hence, 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 extraordinarily 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/ 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. Other patterns, however, have other limits. The threshold for detecting line disparity is about 12 arc seconds [Howard and Rogers 1995]. Also, 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 Cagnello 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 because of 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 2 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 prior studies used conventional monitors (1024 $\times$ 768 resolution) and frame-sequential shutter glasses as the display device. Because human stereoscopic depth perception can take into account a very small difference in the images presented, it is possible that they have considerably underestimated the importance of stereopsis in perceiving large structures, as well as the size of the largest structure that can be clearly viewed.

A criticism that was 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.

Ware and Frank was the only study (we are aware of) to systematically vary the size of the graph while investigating path tracing in 3D graphs, but, for all conditions, the number of nodes and the number of links remained in a constant ratio. It seems intuitively likely that the number of links is a more important factor in the difficulty of tracing paths than the number of nodes.

The present study was designed to address all three of these issues: display resolution, layout, and the relative effect of number of nodes and number of links. 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 antialiased 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 et al. 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 drawn as 3D tubes with graphs drawn using lines. [Note: we use the word "edges" when we refer to graphs as mathematical abstraction and links when we refer to rendered diagrams].

In our first experiment, we varied the viewing method and the sizes of the graphs. Our second experiment was designed to address the importance of the number of links in determining the size of the graph that can be viewed in 3D.

3. EXPERIMENT 1: SPRING LAYOUT GRAPHS WITH AND WITHOUT STEREO AND MOTION CUES

Experiment 1 was close to being a replication of Ware and Frank [1996], but done using spring layout graphs and our ultrahigh-resolution stereoscope. The key independent variables were whether or not stereo or kinetic depth information was available and the size of the graph. We used considerably larger graphs than those investigated in the previous study, because it was clear from pilot work that much larger graphs would be perceivable. An additional independent variable in the present study was the rendering style; displaying the links using solid 3D tubes was compared to using simple unshaded lines. As with previous work, the main dependent variable was the error rate, although time to respond was also measured.

3.1 Task

In each trial, the subject was presented with a graph having 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 (a forced-choice response). Each viewing condition was displayed for a maximum of 5 s, 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. Trials were given in blocks using the same graph with the same layout, but different pairs of highlighted nodes on each individual trial. Our dependent measures were error rate and time to respond.

3.2 The Display

Our display is illustrated in Figures 2 and 3. It consists of a Wheatstone mirror stereoscope [Wheatstone 1838]. This use of front surface mirrors has the advantage that there is no ghosting or reduction in image brightness, problems that plague many stereoscopic display technologies. The displays were Viewsonic VP 2290b monitors. Each of these displays has 3840×2400 pixels, with a display area of 47.7×29.7 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 approximately 24 s 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 one-half of each display.

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, nonstereo perspective projection of the 3D graph.
- **Stereo.** Participants viewed the graphs as a stereo pair generated with an assumed eye separation of 6.4 cm.
- **Motion.** Participants viewed the graph rotating smoothly at a rate of one complete cycle every 36 s.
- **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.

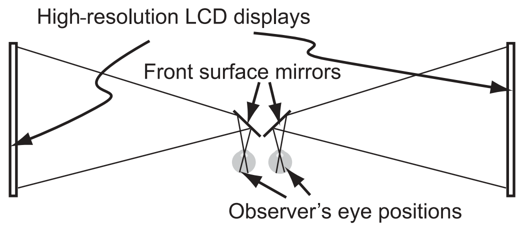


Fig. 2. A Wheatstone stereoscope arrangement provides a ghost-free display.

Fig. 3. A subject viewing the display.

3.4 The Graphs

The algorithm randomly assigned links in such a way that the following statistics resulted (rounded to the nearest integer percentage): 6% of the nodes had degree one (leaf nodes); 37% had degree two; 45% had degree three; 10% had degree four; 2.0% had degree five or greater.

The graphs were laid out using spring forces iteratively applied [di Battista et al. 1999]. There were three kinds of forces used in the layout. (1) Nodes repelled each other with a force inversely proportional to the square of the distance between them. (2) 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). (3) In order to make a more compact overall structure, nodes were subjected to independent forces along the x, y, and z axes toward 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 layout, we used a 3D grid extending the method reported by Fruchterman and Reingold [1991]. By overlaying a $5 \times 5 \times 5$ 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 (nonneighboring) 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.

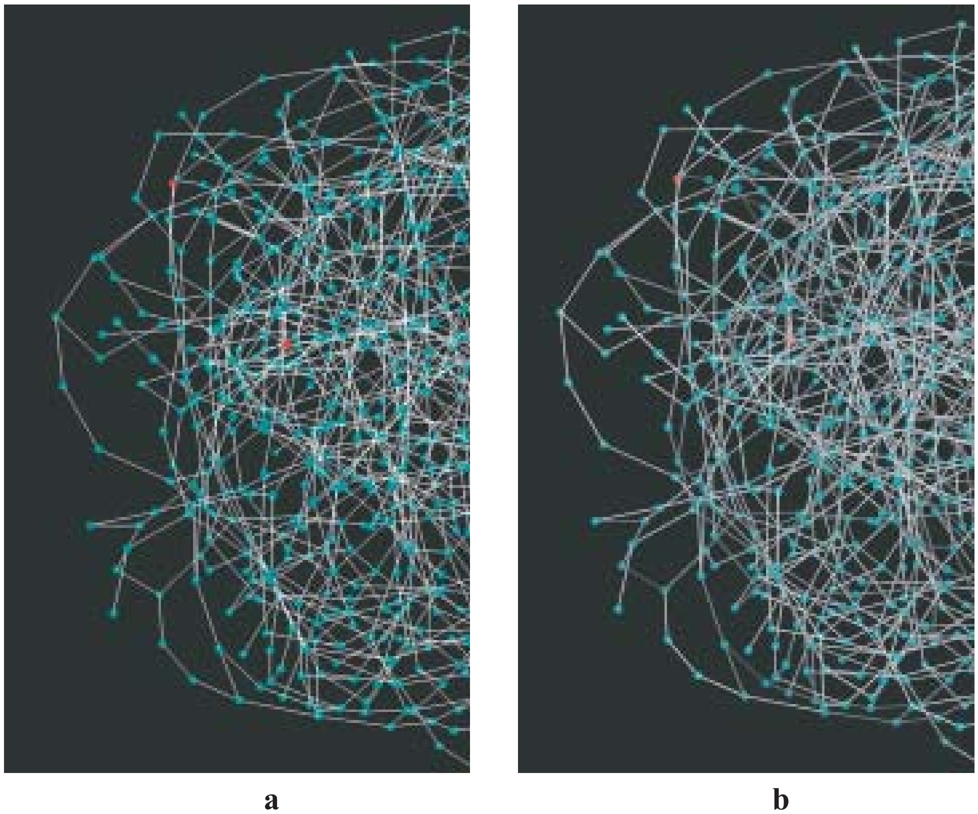


Fig. 4. One-half of a 1000-node graph drawn (a) with lines for the links and (b) with solid tubes for the links.

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 in space 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 confounding variable, our software selected paths in such a way that the mean Euclidean distance between start and end nodes was the same.

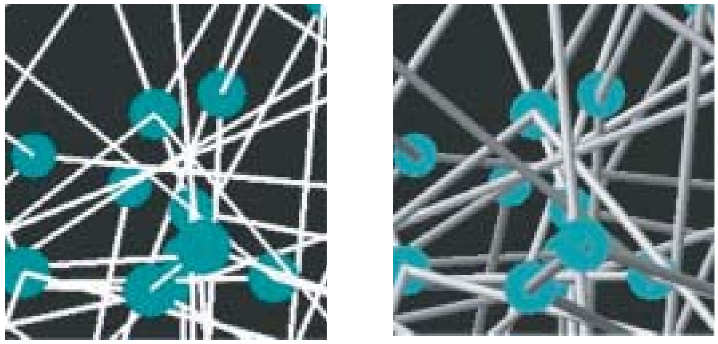
All graphs were rendered against a black background. Line links were drawn on a line thickness of 3 pixels. Cylinder links had a radius of 0.036 cm. The node diameter was 0.3 cm. In the motion conditions, the entire graph rotated about the y axis at a rate of 0.5° per frame (20 fps = 10°/s). Highlighted 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



a b

Fig. 5. Snippets of the 1000-node graph roughly 2× actual size. (a) Line rendering; (b) Tube rendering. Should be viewed from 2 m to get equivalent visual effect.

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 five viewing conditions with both small- and large-sized graphs. A within-participants design was used with each subject exposed to every condition. The experiment took a little over an hour, with at least two short breaks included.

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 and 3D layout without 3D depth cues, except in the middle range of 100 to 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 ($F(3, 39) = 21.6; p < 0.001$) and a main effect for condition ($F(4, 52) = 43.2; p < 0.001$). There was also significant effect for whether tube 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.

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

Figure 8 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 versus 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 plus motion conditions and the other contained the non stereo viewing conditions. Average response times were 15% faster with stereo than without (2.1 s versus 2.42 s).

Figure 9 summarizes the response time data for the experienced observers. Overall our responses took longer than the inexperienced observers, especially with the larger graphs. Response times were also the quickest for the stereo condition; stereo plus motion condition responses were slower.

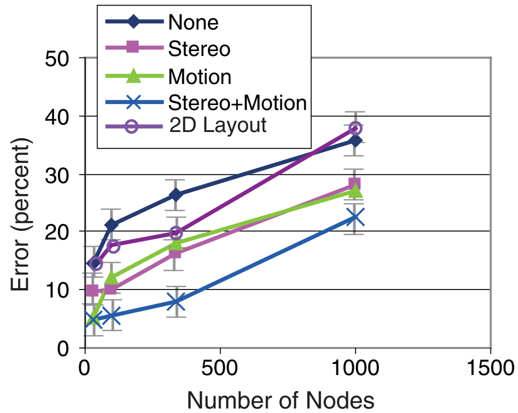


Fig. 6. Errors as a function of graph size. Averaged data from 14 inexperienced participants. Standard error bars represent intersubject variation.

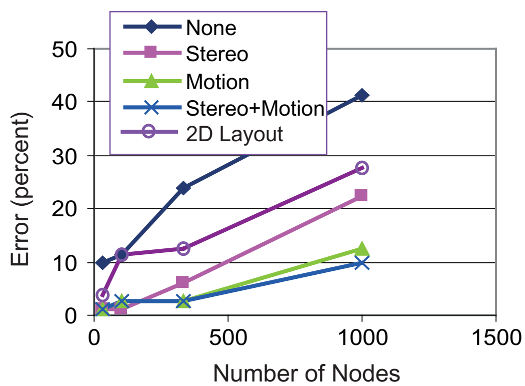


Fig. 7. Errors as a function of graph size. Averaged data from 2 experienced participants.

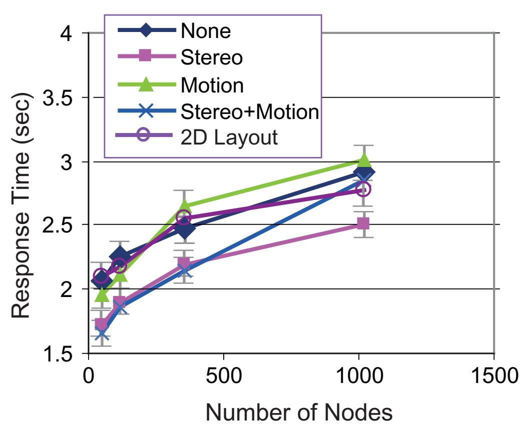


Fig. 8. Time to respond as a function of graph size for the different conditions. Averaged data from 14 inexperienced participants. Standard error bars represent intersubject variation.

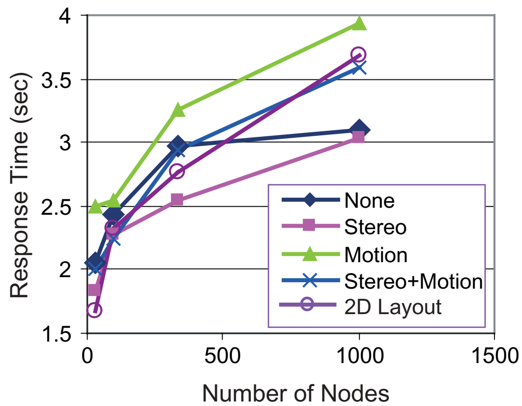


Fig. 9. Time to respond as a function of graph size. Averaged data from 2 experienced participants.

5. DISCUSSION OF EXPERIMENT 1

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 graph 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 threefold 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.

6. EXPERIMENT 2: VARYING THE NUMBER OF NODES AND THE NUMBER OF LINKS

Besides 3D viewing condition and graph size, there are many other factors that may effect our ability to visually trace out paths. In particular, the number of links in a graph may be as or more important than the number of nodes. Alternatively, the ratio of nodes to links could be the critical variable in determining how large a graph can be clearly viewed in 3D.

With our second experiment, we varied both the number of nodes and the number of links in the hope of finding some simple function relating these variables to perceptual traceability for 3D graphs.

6.1 Method

We varied the number of nodes and the number of links in a 3×3 design yielding nine different graph sizes. A pilot study suggested that the ratio of links to nodes (as opposed to the absolute number of links) might be the most critical variable in determining the visual traceability of short paths.

To achieve different edge to node ratios we developed the following random process.

Algorithm
n is the number of nodes
rand() returns a random number between zero and
  one.
dfactor is used to control the edge/node ratio.

```
for(i=0 to n-1)
{
    m = 1
    while (rand() < dfactor)
       m = m+1;
   for (j = 0 to m-1) // add m edges to ith node
    {
       insert edge (i,k).
       k is randomly selected
       from the set of node excluding values
       where (i=k) and (i,k) is already in the graph.
    }
}
```

The *dfactor* values used were: 0.15, 0.295, and 0.4. These values were determined by trial and error to produce, then set, edge/node ratios used for the study.

The nine conditions that resulted were the product of the following node cardinalities and edge ratios.

Nodes:	300, 600, 1200
Edge ratios:	1.17, 1.43, 1.66

6.2 Layout

We used the same spring layout algorithm as for Experiment 1 with two modifications. Because more highly connected graphs resulted in a more compact 3D layout, we added a parameter that scaled the graph about its center, after layout to fill an equivalent 3D volume. In addition, in order to make more effective use of the screen space, the graphs were scaled more in the horizontal direction than the vertical direction.

We reduced the node diameter to 0.2 cm and gave the links a width of 2 pixels (approximately, 0.0123 cm).

6.3 Participants

The participants were 11 graduate and undergraduate students paid for participating.

6.4 Results

An analysis of variance was run on the factors of number of nodes and ratio of links to nodes. There were main effects for: number of nodes ($F(2, 10) = 77.7$, $p < 0.001$), link ratio (($F(2, 10) = 105.1$, $p < 0.001$), and an interaction between nodes and link–node ratio ($F(4, 40) = 3$, $p < 0.05$).

To find out if the ratio between nodes and links was the important variable in determining graph readability we plotted the error rate against this ratio (see Figure 10). If this ratio were a good predictor of error rate, then the points should be approximated a single straight line. Evidently, this is not the case. The curves for the different graph sizes (defined by number of nodes) are offset vertically from one another.

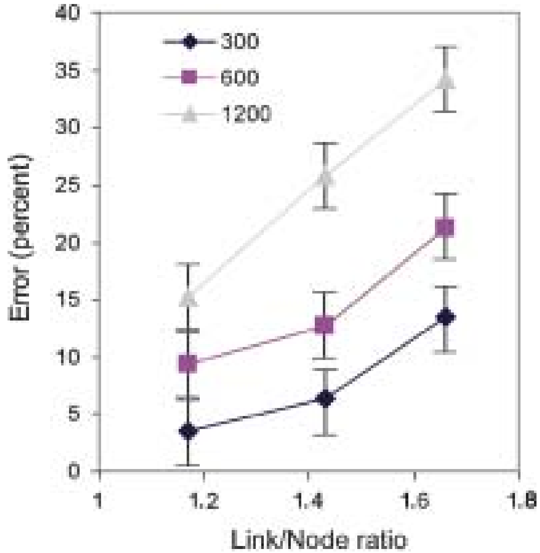


Fig. 10. Error rate is plotted against links/nodes ratio for the three sizes of graphs.

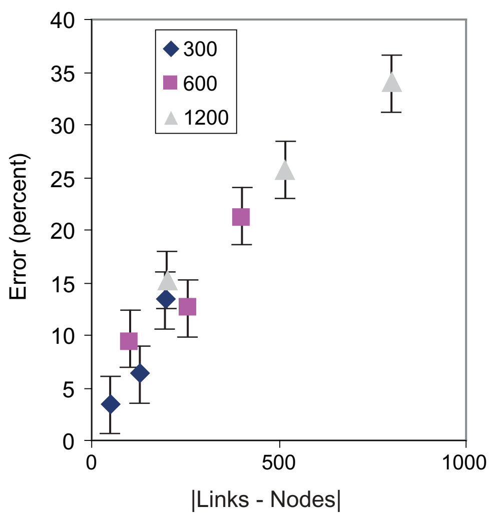


Fig. 11. Mean error rate is plotted against the difference between the number of links and the number of nodes.

After some exploratory data analysis, we discovered that the error rate data could be described much better by the *difference* between the number of nodes and number of links. Figure 11 is a plot of error rate against the number of nodes subtracted from the number of links. As can be seen, the result approximates a straight line. If n is the number of nodes and m is the number of links, computing a linear regression through the points yields

$$error = 4 + 0.04(m - n) \tag{1}$$

with a Pearson r^2 value of 95%.

6.5 Discussion of Experiment 2

The results clearly support our conjecture that the number of links is as important in determining the readability of a graph as the number of nodes. They showed, however, that it is not the ratio between the number of links and the number of nodes, but the difference that determines the error rate. This was wholly unexpected.

The degree of correlation we found was surprisingly high and the regression equation suggests that adding 25 more links for a given number of nodes results in an additional 1% increase in error rate. However, we cannot assume that the size of the graph that can be viewed is without limit. The straight line relationship cannot be expected to hold for very high and very low error rates. It is more likely that what we observed was the central sections of a set of S-shaped curves. Small (or negative) link–node differences would produce low errors, but not negative errors. Similarly, the curves would be expected to flatten out as the errors approached a chance level of 50%.

7. GENERAL DISCUSSION

We have shown that node–link diagrams containing between 500 and 1000 nodes can be accurately "read" in the sense that short paths can be traced out. These graphs are several times larger than previously reported as readable. For comparison, Ware and Frank [1996] reported error rates of about 12% for a 111-node, 148-link graph viewed in stereo with motion parallax. According to the regression equation derived from Experiment 2, this should have resulted in only 5.5% errors using our layout and display system.

There are a number of factors that could account for the discrepancy with prior work, aside from the increase in screen resolution. In the previous work, the graphs were randomly laid out and consisted of a large number of small connected components. The diameters of the nodes and the thickness of the links were larger. The update rates also differed, however, the present experiment actually had lower update rates (20 Hz) than the previous one (30 Hz). Finally, there is the difference between frame sequential stereoscopic display and the use of a mirror-based Wheatstone stereoscope. Frame sequential displays usually have some ghosting, particularly for bright lines on a dark background. A mirror stereoscope has no ghosting at all. Further studies will be required to address the issue of the relative importance of layout, resolution, and stereo display method on large graph perception.

We wish to add a phenomenological observation to our results. We find the high-resolution display more pleasant to view than low-resolution displays using shutter glasses. It is possible that aliasing effects and ghosting may be factors that contribute to the eye strain often reported with stereoscopic viewing systems, although the drawback of a mirror stereoscope is that the head must be held in a constant position. We make no claims regarding the usability of our particular high-resolution stereoscope as a tool for practical examination of large graph structures. However, we do think that it is important to work with the highest quality stereo displays, because we want our results to stand the test of time. We take it as a given that technology continuously improves and that, at some time in the future, high-resolution stereo displays will be widely available if it can be shown that they confer a clear benefit.

Concerning the practical value of 3D display of graphs, the sheer size of the benefit suggests that there may be value in displaying social nets or communication nets in this way. We make no claims regarding data that is not representable in the form of node–link diagrams, but node–link diagrams are very widely used in practice. Therefore, we have hopes that 3D display may become a useful tool in data analysis.

One of the anonymous reviewers of the original submission of this paper pointed out that there are better layout algorithms available [e.g. Huang et al. 1998; Kleiberg et al. 1998] and so our results may not generalize. We do not agree. Although the layout quality will undoubtedly have some impact, we believe that the effect is likely to be small. Random graphs, lacking structure, cannot yield a well-structured layout. We believe, in any case, that the main factor causing errors in tracing paths is the overall density of the edges and this would be unlikely to be much altered. Also, better layout should only increase the size of graphs that can be traced when viewed in 3D.

Probably the most interesting finding we obtained from a theoretical point of view is the simple relationship between number of nodes, number of links, and errors. What we have discovered is analogous to a well known theoretical result. Random graphs, with a node to link ratio of $< 1/2$ have a very high probability of being planar [Bollobas 2001]. This is called a "phase change." This is a theoretical result, however, not necessarily relating to the value of a particular layout in visual thinking, although planar graphs laid out without edge crossings are generally assumed to be easy to perceptually "read."

Figure 11 (and Eq. 1) suggest that there may be a kind of perceptual "phase change" that occurs for random graphs viewed in 3D. For a graph visualization to provide useable support for reasoning task that involves tracing links between nodes, a low error rate is important. To a rough first approximation, we may achieve error rates of 4% (which is likely all that may be expected for rapid responses) when the number of links is no greater than the number of nodes. Of course, it is essential that motion parallax is used to make this possible. This result is potentially useful in that it suggests a target for graph simplification. As with all psychological phenomena, the relationship does not have the precision of a mathematical formula, but does suggest a way of easily expressing those graphs that can be visually traced with low errors. It will be very interesting to discover if a similar relationship holds for other viewing conditions and graphs with other characteristics.

ACKNOWLEDGMENTS

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 antialiasing. Rusty Bobrow of BBN Technologies provided support throughout.

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.

Bollobas, B. 2001. Random Graphs, 2nd Ed., Theorem 6.19, Cambridge Univ. Press, 151.

Campbell, F. W. and Green, D. G. 1965. Monocular versus binocular visual acuity. *Nature, London 208*, 191–192.

di Battista, G., Eades, P., Tamassia, R., and Tollis, I. G. 1999. *Graph Drawing: Algorithms for the Visualization of Graphs*. Prentice-Hall, Englewood Cliffs, NJ.

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.

Fruchterman, T. and 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, M. L., Eades, P., and Wang, J. 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, H., and van Wijk, J. J. 2001. Botanical visualization of huge hierarchies. *IEEE Symposium on Information Visualization*, San Diego. 87–94.

Munzner, T. 1997. H3: Laying out large directed graphs in 3D hyperbolic space. In *Proc. IEEE Symposium on Information Visualization*. 2–10.

MUNZNER, T., GUIMBRETIÈRE, F., AND ROBERTSON, G. 1999. Constellation: A visualization tool for linguistic queries for MindNet. In *Proc. IEEE Symposium on Information Visualization* (Oct.). San Francisco, CA. 132–135.

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., AND BEDERSON, B. B. 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, London* 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.

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

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

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.

WARE, C., ARTHUR, K., AND BOOTH, K. S. 1993. Fishtank virtual reality. *INTERCHI'93 Proceedings.* 37–42.

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

Received March 2006; revised August and December 2006; accepted January 2007.