Personal and Ubiquitous Computing (2013) 17:93 - 105
DOI 10.1007/s00779-011-0473-2

Establishing Aesthetics based on Human Graph Reading Behavior

Weidong Huang

Received: 15 February 2011 / Accepted: 17 August 2011 / Published online: 29 October 2011

Abstract A great deal of real-world data have graph structures and such structures are often visualized into node-link diagrams for a better understanding of the data. Aesthetic criteria have been used as quality measures to evaluate effectiveness of graph visualizations in conveying the embedded information to end users. However, commonly applied aesthetics are originally proposed based on common senses and personal intuitions; thus, their relevance to effectiveness is not guaranteed. It has been agreed that aesthetics should be established based on empirical evidence and derived from theories of how people read graphs. As the first step to this end, we have conducted two eye tracking studies in an attempt to understand the underlying mechanism of edge crossings, the most discussed aesthetic, affecting human graph reading performance. These studies lead to the findings of an important aesthetic of crossing angles and a graph reading behavior of geodesic path tendency. We demonstrate that eye tracking is an effective method for gaining insights on how people read graphs and that how aesthetics can be established based on human graph reading behavior.

Keywords Graph visualization · Graph comprehension · Aesthetics · Edge crossings · Crossing angles · Geodesic path tendency · Eye tracking

1 Introduction

A large portion of real-world data sets have relational structures consisting of a set of entities and relationships between these entities. Examples of such data include physical computer networks, social networks, protein-protein interaction (PPI) networks and World Wide Web. Graphs, defined as a set of vertices and a set of edges that connect the vertices, have been widely used to model such structures for analyzing and understanding the data in consideration [5]. Vertices are used to represent entities while edges are used to represent the relationships. Graphs are often drawn or visualized as node-link diagrams for a better understanding of the underlying data. Figure 1 shows a diagram representing an advice network formed by an auditing team [21], in which ellipses represent managers, diamonds represent staff auditors and boxes represent secretaries. A line from Donna to Nancy, for example, indicates that Donna seeks advice from Nancy.

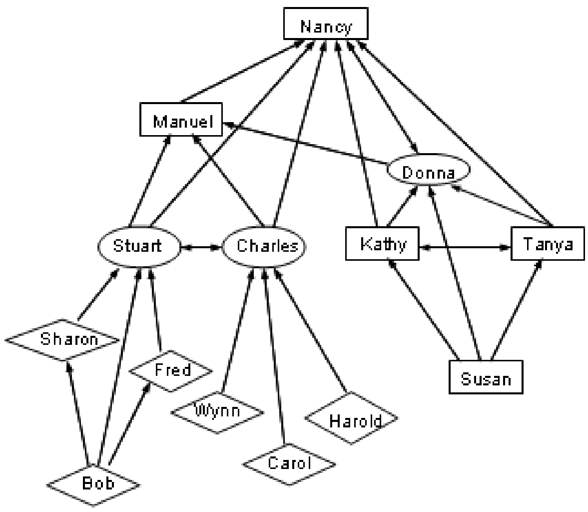


Fig. 1 A node-link-diagram representation of an advice network (reproduced from Krackhardt [21]).

Visualizations of graphs are only useful if the associated node-link diagrams are readable, that is, can effec-

Weidong Huang
University of Tasmania
Newnham, TAS 7248, Australia
E-mail: Tony.Huang@utas.edu.au

tively convey the embedded information to the viewers. When it comes to graph visualization, the first issue we face is how to lay out nodes and links. Given the same graph, it can be drawn in indefinite ways by simply changing the positions of nodes. Research has shown that spatial layout has an effect on human graph comprehension (e.g., [14,23,29]). Good layouts help people to perceive the embedded information quickly and correctly, while poor layouts may confuse people, even convey misleading information. For example, McGrath et al. [23] conducted a user study, in which the users were asked to perceive five different drawings of the same network and answer questions about the structural features of the network. It was found that the perceptions of the users changed while the layout changed.

A range of aesthetic criteria (rules for laying out graphs) have been proposed by researchers from the graph drawing community. Examples of commonly applied aesthetics include minimization of edge crossings, even distribution of vertices and display of maximum symmetries. Graph drawing research concerns the problem of constructing geometric representations of abstract graphs (graphs without specific domain information attached). That is to design algorithms that take a set of vertices and edges of a graph as an input, calculates the positions of the vertices to optimize a set of pre-specified aesthetics. It is assumed that drawings conforming to those criteria are more effective. The past two decades have seen a fast-growing body of research dedicated to constructing algorithms based on aesthetics, in an attempt to produce visually pleasing and easy-to-read graph drawings. For an excellent review on the research of graph drawing, see [5].

However, there are two issues with the current approach of drawing graphs according to pre-specified aesthetics. The first is that commonly applied criteria were originally proposed based on common senses and personal intuitions of researchers; their relevance to human performance in perceiving the embedded information is not guaranteed [29]. In other words, little is known whether or how well the aesthetics actually help to convey the embedded information to the viewer. Therefore, evaluations with real users are often required to ensure effectiveness after the visualization had been done. In evaluating graph visualizations, performance measures such as time and error are widely used to measure effectiveness (e.g., [12,33]). However, these measures treat human perception and cognition as a "black box". They only tell us *what* the effects are when a particular layout is used, but cannot tell us *how* and *why* those effects happened, leaving us wondering where (which parts of the visualization) the time was spent and how the error was caused. As a result, knowledge gained from these measures is useful only for a specific visualization method or task. This knowledge is unlikely to be generally useful when the method or task is varied.

The second is that the criteria used to draw graphs so far have not been specifically related to people's graph reading behavior. The end result of a visualization process is that people read the graph and understand the data. To make a truly effective visualization, it is essential to construct the visualization based on how people read graphs [13]. However, despite the increasing popularity of graphic communication, relatively little is known about how people actually extract and process information from graphs [17,18]. For automatic graph drawing, the ideal approach would be designing algorithms according to aesthetic criteria that are derived from theories of how people read graphs. In this approach, although some human tests will still be necessary, since the criteria used to draw graphs are directly made based on human graph reading behavior, the resultant visualizations will have higher possibility of being more effective.

The work presented in this paper makes an initial attempt toward the understanding of how people read graphs. We conducted two eye tracking studies in which both task performance and eye movements were recorded and analyzed in relation to each other. These two studies resulted in the findings of an important aesthetic of crossing angles and a graph reading behavior of geodesic path tendency. We demonstrated that eye tracking, when used appropriately, can be an effective method in gaining insights on how people read graphs and that how aesthetics can be derived based on human graph reading behavior.

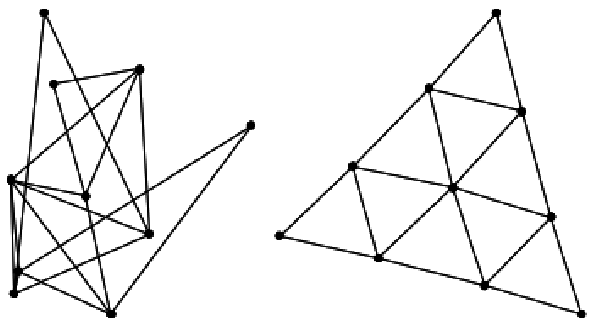


Fig. 2 Two drawings of the same graph. The drawing with crossings on the left fails to display the structure information.

To be more specific, the aesthetic of *edge crossings* has been shown to have the greatest effect on the readability of node-link diagrams [27]. Figure 2 gives a simple demonstration of its effect. Two drawings in this figure are of the same graph. It can be seen that when this graph is drawn with crossings, the structural feature of being symmetric is hidden. Edge crossings has been one of the most discussed aesthetics in graph drawing research and much effort has been devoted to minimizing the number of crossings (e.g., [5]). However, empirical studies in the literature have shown that in some situations, edge crossings may not be as bad as we normally think (for more details, see section 2.2). This means that there is a need for understanding of the underlying mechanism of edge crossings affecting human performance. Therefore, as a starting point, we began our process of understanding how people read graphs by addressing this need.

The remainder of this paper is organized as follows. In section 2, we review the literature with a focus on eye tracking studies of reading diagrams and on empirical studies of edge crossings. Then our two eye tracking studies are presented in section 3 and section 4. Finally in section 5 we conclude the paper with a general discussion.

2 Related Work

With the advance of technology and computer hardware, the use of eye tracking has become increasingly more affordable and popular in recent years. Although eye tracking has a long history of being used for general diagrams such as maps, statistic graphs and scientific visualizations in the fields of Psychology and Education [32], as far as we know, eye tracking has not been used specifically in relation to aesthetics of node-link diagrams. In this section, we briefly review these two separate bodies of research. The reviewed research also serves as a solid foundation for our work presented in this paper.

2.1 Eye Tracking Studies of Diagrams

There is a fair amount of eye tracking research available in the literature which investigates how people retrieve and process information from diagrams. For example, Carpenter and Shah [3] recorded eye movement data while subjects answered questions in order to examine cognitive processes involved in the comprehension of line graphs. Lohse [22] used eye movement data to examine how individual differences in memory capacity and changes in graphic design can affect graphical information processing. Peebles and Cheng [25] examined eye movement patterns to construct and test cognitive models of graph reasoning for a common task of eliciting the value of one variable corresponding to a given value of another. Ratwani et al. [30] used eye tracking in a study and found that people employed different sets of cognitive processes to extract local and global information from so-called choropleth graphs (a type of spatial color-coded diagrams). Convertino et al. [4] conducted an eye tracking study in order to investigate how users integrate data from multiple diagrams.

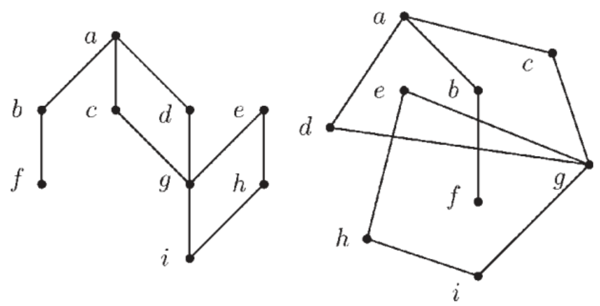


Fig. 3 Two examples of the hierarchical graphs used in the experiment of Korner [18].

Although eye tracking has been widely used to investigate eye movement patterns and corresponding cognitive processes, the literature related to how people read graphs represented as node-link diagrams is sparse. One notable exception is the pioneering study of Korner [18]. This study used eye tracking to test two possible models of comprehension of hierarchical graphs. Eight drawings that represented the same information of preference and varied in spatial layout were given; see Figure 3 for examples. Subjects were asked to give yes or no responses to simple questions such as "Is a better than b?" Eye movements were recorded during task performance. The detailed analysis of eye tracking data suggested a three-stage model of graph comprehension: two stages searching for the target nodes, followed by a separated reasoning stage. This model was further tested and validated in a recent study by Korner [19], suggesting that people read graphs in a sequential manner. Korner further mentioned that this manner is an economical approach of comprehension which reduces the overall processing load. Pohl et al. [26] conducted an eye tracking study to compare the readability of different graph layouts. In this study, eye movement data (heatmaps) were used to analyze the graph reading strategies of the subjects in order to explain why one layout was better than another for a particular task.

2.2 Aesthetic of Edge Crossings

In drawing graphs as node-link diagrams, it is commonly accepted and employed as a general rule that the number of edge crossings should be minimized whenever possible. A number of user studies have been conducted to examine the effect of edge crossings in a variety of experimental settings (e.g., [14, 20, 27]). However, as reviewed below, research in the literature has so far presented a mixed picture on what had made crossings important, highlighting the need of understanding the processes that underlay the observed effects.

Purchase et al. [29] conducted a user study and provided the first empirical evidence that validates the aesthetic of edge crossings for its negative impact on human comprehension of graphs. In another study, Purchase [27] found that the aesthetic of crossings was the most important factor affecting graph reading performance, compared to other four aesthetic principles. Further, Ware et al. [33] used abstracted graphs and asked subjects to perform the shortest path search task. They found that it was the number of crossings on the shortest path that significantly affected performance, rather than the total number of crossings. Korner and Albert [20] conducted a study in which hierarchical graphs of ordered sets were used and three visual properties were compared: planarity (edge crossings), slopes and levels. It was also found that planarity was the most influential factor affecting response time. This study further revealed that “it is the general disarrangement present in crossed drawings that causes the slower comprehension speed”. This means that even though the lines associated with the task do not cross any other edges, the performance can still be affected by the crossings in the drawing. Huang et al. [14] conducted a study investigating which layout is suitable for which task and found that in perceiving sociograms (node-link diagrams for social networks), crossings are important only for tasks that involve path tracing.

In addition, some researchers have pointed out that crossings with different crossing styles may have different degrees of impact. For example, as suggested by Ware et al. [33], crossings with nearly-90-degree angles can be less confusing than those with acute angles. Further, more and more empirical studies are available showing that in some situations, crossings may not be as bad as we normally think. For example, when sociograms are drawn to convey information about groups, it may be more desirable to cross edges connecting the group members [15]. Crossings are also necessary when we want to display important structural features, such as symmetry; drawing graph without crossings may hide some dimensions of symmetry.

3 Experiment One

This experiment was largely exploratory and our aim was to gain initial insights on the underlying mechanism of crossings affecting human graph reading by observing their eye movements.

3.1 Design

A within-subject design was employed. We used social networks as the experimental data. Each network was drawn twice with one having no crossings on the shortest path of two pre-specified nodes (there could be crossings on other parts of the drawing), and the other having a few crossings. Subjects were asked to answer questions specific to the network in consideration. Their performance data were logged in real time by a custom-built system and their eye movements were video-recorded by an eye tracker. We analyzed the performance data in relation to the eye movements to see how the performance was affected. Questionnaires and interviews were also conducted in order to gain further insights on the eye movement behavior of the subjects.

It should be noted that traditional eye movement measures such as the number of fixations and the mean of fixation durations are often used in eye tracking studies for analysis of cognitive processes [10, 32]. These measures are useful for finding collective eye movement patterns for quantitative analysis, and those patterns are often related to different areas of interest in the stimulus or different time periods of the whole cognitive process. However, it is difficult to relate these eye movement measures to specific graph elements such as nodes, edges and paths. Given the qualitative nature of our purposes for this study, we instead used eye movement videos for the analysis. Eye movement videos gave us an immediate continuous view of how visual queries were executed and which node or edge the viewer was looking at during the process of graph comprehension.

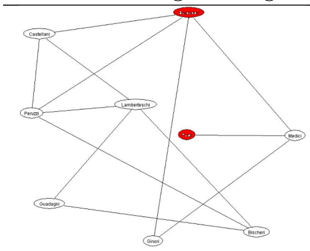
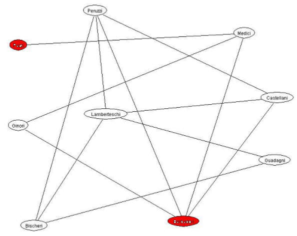
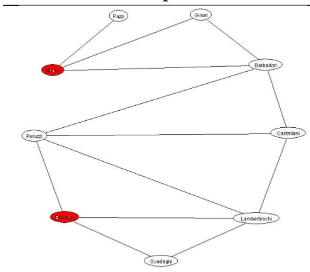
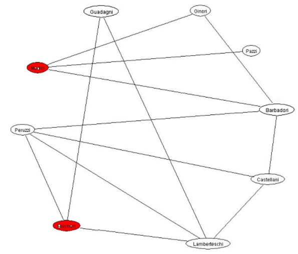
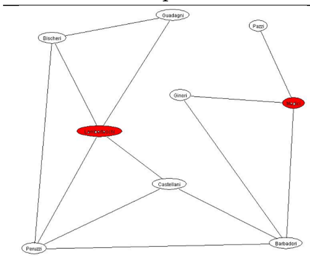
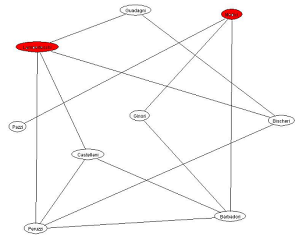
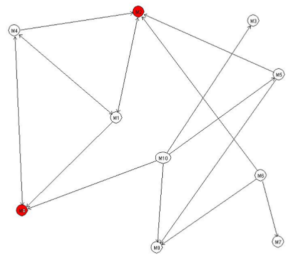
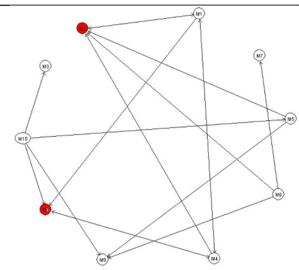
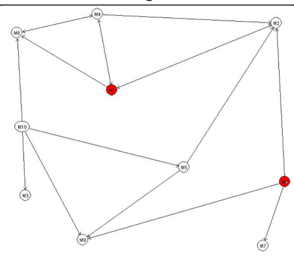
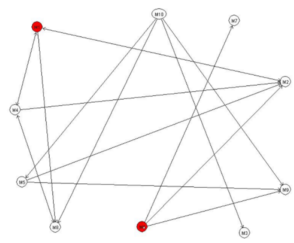
Further, eye movement data only tell us *how*. To gain information about *why*, the method of think aloud may be useful. This method has been successfully used together with eye tracking in studies in which relatively complex pictures and difficult tasks were used (e.g., [30, 31]). However, think aloud involves asking subjects to verbalize whatever he/she is doing and thinking during the task performance, which can be intrusive. Also, in performing simple perception tasks, human eyes move faster than the pace they think [7, p140]. Employing eye tracking and think aloud at the same time in our study may distort eye movement patterns, making them misleading. We therefore conducted post-task interviews instead. That is, ask users to reflect and explain the eye movement behavior while watching their own eye videos.

3.2 Stimuli

Three graphs were obtained based on two social networks of Wasserman and Faust [34]: Padgetts Florentine families business relations and Krackhardts high-tech managers friendship relations. The first graph was the full set of the family business data with 11 nodes and 15 edges. The second was a subset of the same data with 9 nodes and 13 edges. The third was a digraph and a subset of the manager friendship data with 10 nodes and 14 edges.

Ten drawings were drawn by hand from the three graphs and they are shown in Table 1. Note that each pair of the crossing and non-crossing drawings on the same row was of the same graph. The drawings in pairs 4 and 5 were of the digraph and drawn with edges having arrows indicating the direction of relationships between two managers. Two drawings of each pair also had the same pair of nodes highlighted. The highlighted nodes were to be used for questions.

Table 1 Five pairs of drawings used in experiment one.

Non-crossing drawing	Crossing drawing
 1p	 1c
 2p	 2c
 3p	 3c
 4p	 4c
 5p	 5c

3.3 The Task

There was only one question for each drawing. For drawings of family business relations, the question was “What is the separation level between the two highlighted families?” The separation level was defined as the least number of edges between the two families. For those of manager friendship relations, the question was “Do the two highlighted managers have friend’s-friend relationship?” Two managers A and C have friend’s-friend relationship if there is a manager B between them, i.e., either A→B→C or A←B←C.

It should be noted that although the questions were specific network questions, they are essentially the shortest path search task; that is, find the least number of edges between two nodes. This was to make the experiment results comparable to those of others, since the shortest path search task has been widely used in various user studies (e.g., [33]).

3.4 Subjects

Thirteen subjects participated in the study. These subjects were postgraduates with normal vision and were completely new to eye tracking. Three of them were experienced in graph reading, two had no experience at all, and others had limited experience only related to database or/and information visualization courses. They were reimbursed $20 each for their time upon the completion of their tasks.

3.5 Apparatus

The experiment was conducted individually in a quiet room. There were one operator PC on which the eye tracking system was running and one subject laptop on which the diagrams were displayed. There were also tables and adjustable chairs in the room. Adjustments were made to maintain the eyes of the subject at a distance of approximately 50cm from the 14-inch monitor of the laptop. A chin rest was used to reduce head movements.

The eye tracker had a helmet that was to be worn by the subject. An eye camera was attached to the helmet to record the eye movements of the subject. A scan converter was installed on the subject laptop and used to record the content of the screen. The eye tracking system tracked eye movements by observing the position of the pupil and corneal reflex from the right eye. These eye positions indicated by a gaze cursor combined with the video signals from the scan converter were recorded in real time into MPEG videos for off-line analysis.

3.6 Online System

The stimuli were displayed on the subject laptop by a custom-built experimental system. The system displayed a question first. The subject pressed the button on the screen, the question disappeared and the corresponding drawing was then shown. The subject answered the question by clicking one of the buttons above the drawing; each button showed one possible answer. Once the button was clicked, a new question was shown and so on. We displayed each question and its corresponding drawing separately to make sure that the response time recorded did not include the time for reading questions.

The response time of the subject for each drawing (which started once the diagram was completely displayed and ended once a button was clicked) and the corresponding answer were recorded by the system in real time.

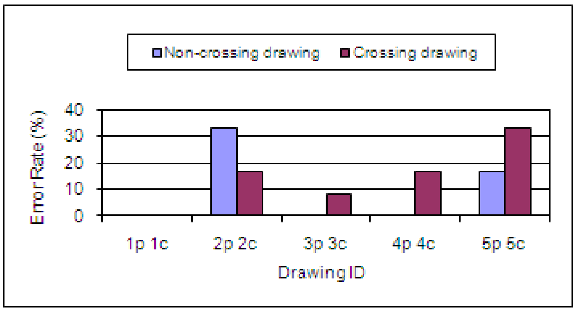
3.7 Procedure

Before starting the experiment, the subject was asked to read through and understand the tutorial materials, and sign the consent form. The subject was also given a chance to practice and ask questions. During the preparation time, the subject was instructed to look for the answer once the drawing was shown, click the corresponding button once the answer was determined and not to look around in between.

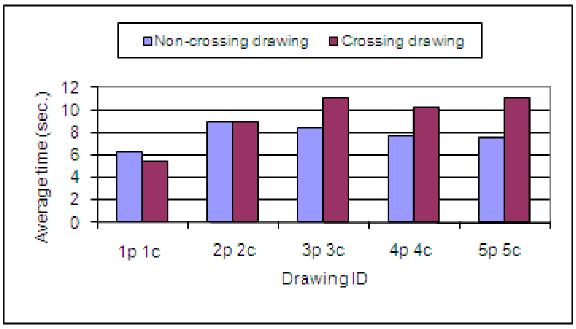
When ready to start, the subject was seated and the helmet was worn. After a short calibration, the online system was started and tasks were performed. The drawings were displayed in a random order. Following the graph reading tasks, a post-task questionnaire and experimental debriefing were given. The questionnaire asked questions about what strategies were used for graph reading; whether edge crossings had impact on their preference and performance; whether there were any other drawing features that could aid or hinder performance, and how those features could affect graph reading behavior. After the questionnaire, the subject was interviewed in relation to their responses to the questionnaire. The subject was also asked to reflect and explain the graph reading strategies he/she took while watching his/her own eye videos. The whole session took about 60 minutes.

3.8 Results

One subject read graphs with the aid of the mouse cursor throughout the testing, which made his testing setting different from others. Therefore, his data were omitted and analysis was based on the data produced by the remaining twelve subjects. The error rates and time averages are shown in Figure 4.



(a) Error rates(%).



(b) Average times (sec.).

Fig. 4 Performance data.

Given the simplicity of the graphs, no time limits on the task and the background of the subjects, the reason for the errors was most likely because the subjects counted the wrong number of edges or clicked a wrong button on the screen. This made it difficult to exclusively relate the error data to the quality of drawings. We therefore report analysis of time data only here. First we consider overall difference between crossing and non-crossing drawings. The time that the subjects spent on the crossing drawings was 9.36 seconds on average, which was 1.58 seconds slower than the time spent on the non-crossing drawings (7.78 seconds). The paired t test indicated that this difference was significant, $t(11)$=-2.34; p<0.05.

Next consider the differences for each pair of crossing and non-crossing drawings. As can be seen from Figure 4(b), there were three pairs (pairs 3, 4 and 5), for which the subjects spent more time with the crossing drawings than with the corresponding non-crossing drawings. There were two pairs (pairs 1 and 2), for which the subjects spent more time with the non-crossing drawings. However, the paired t tests revealed that only the difference between the drawings of pair 5 (5p and 5c) was statistically significant, $t(11)$=-2.47; p<0.05.

3.9 Video Analysis and Discussion

The time data indicated that on average, the time spent with crossing drawings was significantly longer than that with non-crossing drawings. This was consistent with the general finding that crossings negatively affect performance. However, a different pattern was revealed at the individual level. In particular, for the two drawings of pair 1, on average the subjects took about 0.87 second more time with 1p than they did with 1c. This was a surprise, since 1c had 8 crossings on the shortest path, while 1p had no crossings at all.

After examining the video data, it was found that most of the subjects started to search from the highlighted node on the top or left; this also was the case for drawings 1p and 1c. This may be related to the daily reading habits of the subjects. For 1p, the top-left node was the node with four incident edges; the subjects spent some time following these edges trying to find the path to the target node. On the contrary, the top-left node for 1c had only one incident edge. The subjects simply followed this edge to the next node that happened to be a neighbor of the target node.

However, there was still a question: how can one explain the lack of expected impact of crossings in drawing 1c? Surprisingly, a closer look of the video data revealed that crossings had little impact on eye movements. Eyes moved smoothly without delay when passing through crossings. It appeared that the subjects simply ignored the crossings during the search for the target path. This was also consistent with the comments of the subjects regarding perceived effects of edge crossings. Typical responses among them included "drawings with no crossings were better. Crossing drawings looked more confusing. But crossings did not have much impact on me finding answers". "No effect for me. I could see the paths easily with the crossings". One subject commented slightly differently though: "Yeah, there could be an effect when there were two or more paths. It was not that obvious to see which one is the answer when crossings were there. But when it came to my eye movements, I don't think those crossings affected me". After examining the features of our stimuli, we found two possible explanations for the lack of the crossing effects:

1. The graphs used were very sparse and small with the largest graph containing only 11 nodes and 15 edges. The effect of crossings with this kind of graphs could be too small to observe.
2. Crossing angles in the drawings were relatively large. Most of them were close to 90 degrees. Edge crossings with such large angles might not be as confusing as we expected.

On the one hand, eye movement data showed little impact of edge crossings. On the other hand, the time data indicated that non-crossing drawings were indeed more effective than crossing drawings. Therefore, there must be other factors at work leading to the observed differences between crossing and non-crossing drawings. The video inspections revealed that more path search eye movements were involved with crossing drawings and those extra movements were mostly on edges that were not part of the target path. This was also supported by the user comments. For example, "I tend to think that straight paths are short". "If there were multiple edges, I tried the one close to the end node first. But sometimes, I was led to a dead end and had to try another edge". These comments gave a possible explanation for the observed effect of crossings in response time: more time was spent on the distracting branch edges. As shown in Table 1, there were more branch edges being drawn close toward the target node in the crossing drawings, compared with the non-crossing drawings.

Based on the observations of the video data, all the subjects showed similar high level search strategies, which provided further eye movement evidence to the three-stage graph reading model of Korner [18,19]:

1. Look for the highlighted nodes and determine which one to start with for the answer.

2. Search for possible paths. During this process, some important edges and nodes might be visited repeatedly.
3. Determine and verify the answer.

In summary, in this particular experiment, it was clear that it was not the edge crossings themselves that directly affected performance or eye movements. Rather, it seemed to be other features of the crossing drawings that caused observed effects. For example, large crossing angles and distracting branch edges. To get a clearer picture of how crossings affect performance and eye movements, experiment 2 was conducted, which we present in the next section.

4 Experiment Two

Experiment 1 was set to examine the effect of edge crossings on performance and eye movements. As expected, the performance data indicated that subjects performed better with non-crossing drawings. However, the video analysis revealed that the performance difference was not the result of the crossings themselves. This was surprising. Most graph drawing research shows a strong belief that crossings are the major time consuming factor in graph reading, but the results of experiment 1 suggested that they were not.

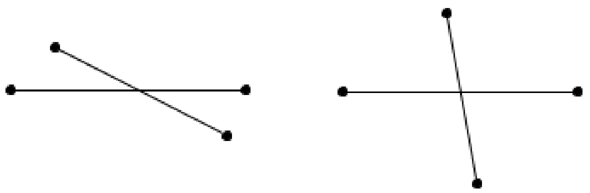


Fig. 5 Illustration of crossing angle effects. The crossing on the right is less confusing than that on the left.

First, it appeared that crossings were simply ignored by the subjects during the path searching; the eye movements remained smooth without apparent delay when passing through crossings. According to Ware et al. [33], we suspected that this could be because the crossing angles were large. As shown in Figure 5, crossings with large angles might not be as confusing as those with small angles.

Second, the extra eye movements were involved with the crossing drawings. Based on the video data and the user comments, we suspected that the subjects tended to follow branch edges going toward the target node, which is illustrated in Figure 6. We term this behavior as geodesic path tendency (geodesic path is the straight line segment between two nodes).

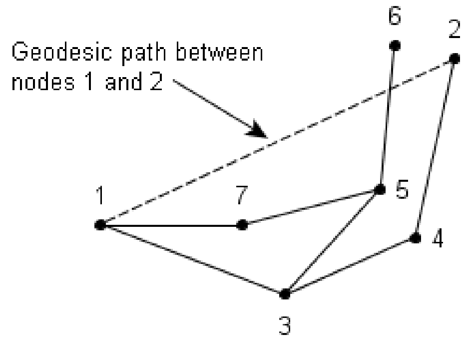


Fig. 6 Illustration of geodesic path tendency. Note that the dashed line is not part of the graph. To find the shortest path between nodes 1 and 2, if search always starts from node 1, people tend to follow the path 1-7-5-6 first, then the path 1-3-5-6, and finally the path 1-3-4-2.

In addition, the fact that the graphs used were too small might also contribute to the surprising results of experiment 1. Experiment 2 was conducted to address the above-mentioned conjectures.

4.1 Design

For this experiment, two sets of drawings were produced using two relatively larger and more complex graphs. We drew one graph with the size of crossing angles being varied and drew another graph with the direction of branch edges being changed in different drawings. We asked subjects to perform graph reading tasks, and their performance and eye movements were recorded. Post-task questionnaires and interviews were conducted.

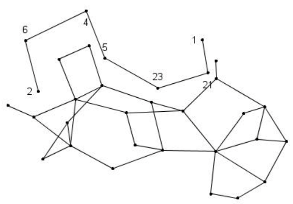
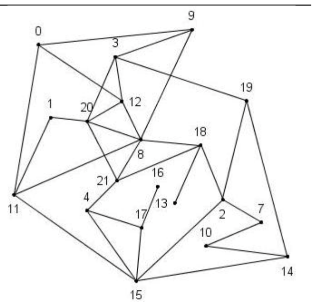
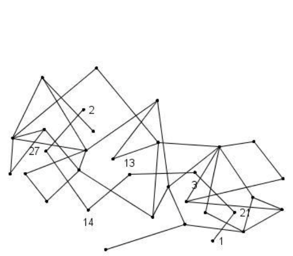
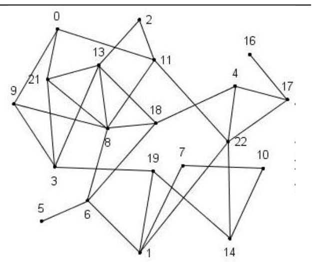
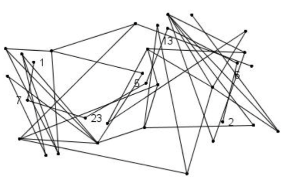
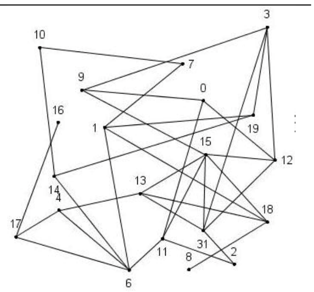
4.2 Subjects

Sixteen subjects were recruited on a completely voluntary basis. All had normal vision and were regular computer users. They had different degrees of familiarity with node-link diagrams; two of them had no knowledge at all at the time of participation. The subjects were reimbursed $20 each for their time upon the completion of their tasks.

4.3 Stimuli

Two sets of drawings were produced based on two graphs. Each set had three different drawings of the same graph. As shown in Table 2, set 1 drawings were for testing crossing angles, and set 2 drawings were for testing geodesic path tendency. Note that for set 1 in Table 2, only some of the nodes were labeled for clarity, though

Table 2 Two sets of drawings used in experiment two.

Set 1	Set 2
 c1	 f1
 c2	 f2
 c3	 f3

in the real tests, all labels were visible. Also note that the same node may have different labels in different drawings.

Set 1: The graph for three drawings (c1, c2 and c3) contained 32 nodes and 43 edges. It had two components: a path component and a condition component. In producing the three drawings, the layout of the path component remained unchanged. The layout of the other component was modified to make three conditions: no crossings on the path (c1), nearly-90-degree crossings on the path (c2) and small-angle crossings on the path (c3). Drawing c1 was the control condition that was to compare how eye movements changed when crossings were introduced in c2 and c3. However, the subjects were not made aware of these facts beforehand.

Set 2: Three drawings (f1, f2 and f3) were of another graph containing 20 nodes and 32 edges. In f1, the shortest path between nodes 1 and 2 (1-11-15-2) was far away from the geodesic path of nodes 1 and 2 and had no crossings, while the shortest path in f2 (1-22-11-2) and f3 (1-6-11-2) had three crossings (with nearly-90-degree angles) and was near the geodesic path. In addition, the total number of crossings was 3 in f1, 12 in f2 and 37 in f3.

4.4 The Task

The task was to find the shortest path between nodes 1 and 2. This shortest path task has been used in two ways in the literature. First, in studies such as the study of Ware et al. [33], the target nodes are pre-specified and highlighted. Therefore, locating the nodes was no longer necessary and the task required only path searching. Second, the nodes in consideration were not highlighted in some studies such as that of Korner and Albert [20].

As mentioned in section 2, while Ware et al. found that it was only crossings on the shortest path that were important, Korner and Albert found that it was the overall arrangement of crossings that mattered. Korner and Albert mentioned that "crossings themselves may affect early stages of visual information processing", and that "such salient properties (crossings) are processed in precedence, and draw attention and distract the visual system from the message of the drawing". If this was the case, then according to the three-stage model of Korner [18], these distractions likely happened during the first two node locating stages. To see what would happen when node locating and path search were used independently and in combination, the following three specific tasks were included:

1. Node task: find the most connected node (the node that has most incident edges).
2. Path task: find the shortest path between nodes 1 and 2. Nodes 1 and 2 were highlighted.
3. Node+Path task: find the shortest path between nodes 1 and 2. Nodes 1 and 2 were not highlighted.

4.5 Procedure

The same online system and eye tracking system in experiment 1 were used to record the experimental data. This experiment included three sessions: one session for each task. The order of the three tasks was random. In each session, subjects had to perform the task with each of the six drawings, which in turn were displayed in a random order. There was a break between sessions. A calibration was conducted before each session. Subjects were given time to read the tutorial materials, ask questions and practice. They were also instructed to answer each question as quickly as possible without compromising accuracy.

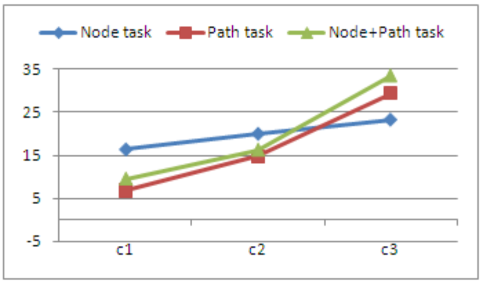
This experiment was conducted on an individual basis. The subjects performed tasks online first. Then a post-task questionnaire was given, followed by a short interview. Seven of the subjects were asked to explain their eye-movement behavior while watching their own eye videos. The whole experiment took about 50 minutes on average.

4.6 Quantitative Results

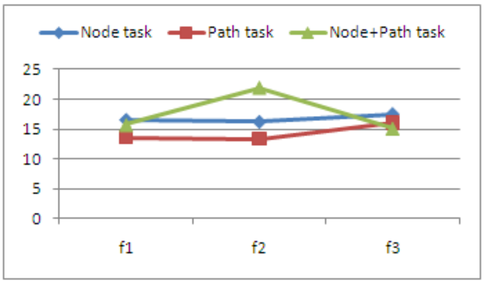
Regarding response time, the averages for all drawings are shown in Table 3 and illustrated in Figure 7. The response time data were analyzed using the non-parametric Friedman tests and post hoc Wilcoxon signed rank tests.

Table 3 Average times (sec.) for set 1 and set 2 drawings.

DrawingID	c1	c2	c3
Node task	16.37	19.97	23.22
Path task	6.81	14.74	29.41
Node+Path task	9.54	16.28	33.58
DrawingID	f1	f2	f3
Node task	16.54	16.20	17.52
Path task	13.61	13.33	16.07
Node+Path task	15.72	21.91	15.12



(a) Average times (sec.) for set 1 drawings.



(b) Average times (sec.) for set 2 drawings.

Fig. 7 Response time data.

Node Task: Among set 1 drawings, the shortest time was spent with c1, followed by c2, then c3. The Friedman test indicated that these differences were statistically significant (χ^2=6.13; p<0.05). Pairwise comparisons found that only the difference between c1 and c3 was significant (p<0.01). For set 2 drawings, the shortest time was spent with f2, followed by f1, then f3. However, the Friedman test did not find any significant differences (χ^2=3.88; p=0.144).

Path Task: For set 1 drawings, the shortest time was spent with c1, followed by c2, then c3. The Friedman test revealed that these differences were statistically significant (χ^2=26.38; p<0.001). Pairwise comparisons indicated that the time difference for each pair was also statistically significant (for each pair, p≤0.001). For set 2 drawings, the shortest time was spent with f2, followed by f1, then f3. However, the Friedman test showed that these differences were not statistically significant (χ^2=1.13; p=0.570).

Node+Path Task: For set 1 drawings, the shortest time was spent with c1, followed by c2, then c3. The Friedman test revealed that these differences were statistically significant (χ^2=18.00; p<0.001). Pairwise comparisons indicated that the difference for each pair was also statistically significant (for each pair, p<0.05). For set 2 drawings, the shortest time was spent with f3, followed by f1, then f2. The Friedman test showed that these differences were statistically significant (χ^2=6.50; p<0.05). Pairwise comparisons found that the differences between f1 and f2, f2 and f3 were statistically significant (p<0.05).

In respect to error rates, the results are shown in Table 4. In particular, for Node task, all the responses for set 1 and set 2 drawings were correct, except for c2 and c3. The error rate was 6.25% for c2 and 25% for c3, respectively.

Table 4 Error rates (%) for set 1 and set 2 drawings.

DrawingID	c1	c2	c3
Node task	0.00	6.25	25.00
Path task	0.00	0.00	0.00
Node+Path task	0.00	0.00	0.00
DrawingID	f1	f2	f3
Node task	0.00	0.00	0.00
Path task	18.75	31.25	0.00
Node+Path task	18.75	37.50	12.50

For the path search tasks (Path task and Node+Path task), regardless of whether the nodes were highlighted or not, the subjects made no errors for set 1 drawings. For set 2, when the nodes were highlighted, the highest error rate was made with f2 (31.25%), followed by f1 (18.75%). No errors were made with f3. When no nodes were highlighted, the highest error rate was with f2 (37.5%), followed by f1 (18.75%), then f3 (12.5%).

4.7 Eye Movement Videos and Questionnaire Analysis

The videos showed clearly that for the path search tasks with set 1 drawings, the speed of eye movements was the fastest with c1. With c2, the overall eye movements were still smooth but became slower. Although some subjects claimed that they were not affected by the crossings here, the response time data did show that the subjects responded significantly slower with c2 than with c1. With c3, eye movements became generally slower; they were slowest on the edge connecting nodes 13 and 6 (see c3 in Table 2). There were also many back-and-forth eye movements around the crossing points on that edge, indicating that the eyes of the viewer were uncertain about which way to go. Clearly, the small-angle crossings in c3 caused slow and extra eye movements, which contributed to the longest response time. The comments of the subjects on crossings included: "(Crossings) forced me to focus harder"; "(Crossings) helped to improve my concentration"; "Crossings affected me except at right angles"; "I think crossings slowed me down"; "Crossings make graphs more complicated and confusing"; "If the angle is small, you have to be careful when following the link to make sure you end up at the right node".

Regarding the path search tasks with set 2 drawings, to find the shortest path in f1, most of the subjects searched on the paths near the geodesic path first. In particular, there were fifteen subjects (94%) for Path task and twelve (75%) for Node+Path task who searched the nearest path of 1-20-8-18-2 first. Some simply missed the target path of 1-11-15-2, which is further away from the geodesic path. The rest of the subjects detected the target path either at a later stage of searching or just before pressing the answer button, as commented by a subject: "I often found the shortest route last".

The high error rates with f2 for the path search tasks were surprising. The video inspection on f2 revealed that the subjects spent most of their time on the left part of the drawing, where there were more crossings (the same also happened on the right part of f3). In addition, most of the subjects found the correct path (1-22-11-2) in f2 at the late stage of the task. In general, six subjects mentioned in the questionnaire that long edges had some influence and commented: "Long edges need more time to reconfirm"; "The shortest path of a few long steps (edges) outside many short steps is harder to see".

In regard to Node task, it appeared that the subjects adopted the same eye movement strategy for this task. That is, eyes stayed around a node for a while counting the number of edges, then moved straight from the node to the next. It appeared that crossings were more or less ignored, while the subjects tended to start the task with nodes in dense areas first. Given this behavior, the significant time difference between c1 and c3 for Node task might be caused by the striking difference in node angular resolution between the two drawings, according to Huang et al. [12]: in c1, neighboring edges were well separated around the nodes, which makes edge counting straightforward, while in c3, neighboring edges were closely attached to the highly connected nodes, making counting difficult. All the subjects claimed that for Node task, crossings did not have any influence on them.

4.8 Comparison of Path Task and Node+Path Task

Path task and Node+Path task were included in the study to see whether and how much the eyes can be distracted by crossings during the node locating stage when no nodes were highlighted. The video analysis showed that eyes appeared not to have been distracted by the crossings. The eye movement patterns for node locating of Node+Path task were very similar to those of Node task, though their eyes moved faster since they only needed to identify the labels, rather than count the number of incident edges. The subjects made the similar comments as they made about the effect of crossings for Node task. That is, there was no much impact on node locating of Node+Path.

Table 5 Performance data for Path task and Node+Path task.

	Time (sec.)	Error rate (%)
Path task	15.77	8.33
Node+Path task	18.69	11.46

The performance data for the path search tasks are shown in Table 5. Regarding response time, the subjects took 15.77 seconds on average to complete Path task, which was 2.92 seconds faster than the time they took to complete Node+Path. The average time for Node+Path was 18.69 seconds. This is normal since the subjects needed extra time to locate the target nodes first. The Wilcoxon signed rank test revealed that this time difference was marginally significant (p=0.046). This means that statistically the extra component of node locating in Node+Path task resulted in only a trivial increase in response time.

With regard to the error rate, the average error rate for Path task was 8.33%, while the average error rate for

Node+Path task was 11.46%. This indicated that when no nodes were highlighted, more errors were made.

4.9 Discussion

The results of this experiment indicated that the eye movements of node locating for both Node task and Node+Path task were largely independent of crossings. This is in good agreement with the finding of Korner [18] that in searching for target nodes, the information provided by crossings was ignored: "After all, the presence or absence of crossed lines does not constrain the position of nodes in the graph".

In this particular study, the extra component of locating nodes in Node+Path task only led to a marginal increase in response time, compared to Path task. This seemed unusual at first. However, it is in fact reasonable since research has shown that the human visual system can be effective in searching for a target object among similar distractors [9,18]. In addition, node locating might have helped the subjects to get ideas about the target path through their peripheral vision. This in turn helped to reduce the time for path searching, thus compensating for the time needed for locating nodes. However, the subjects did make more errors for Node+Path task. One possible explanation could be that when no nodes were highlighted, the target path became less visible, and thus relatively harder for the viewer to detect. This also partially explains the difference between the findings of Korner and Albert and Ware et al.: when nodes were highlighted, subjects were able to focus their attentions on relevant paths, in which only local crossings were involved, making crossings on the shortest path more important. On the other hand, when no nodes were highlighted, the visual aid for node locating was lost. Subjects had to remember or consistently verify the locations of the target nodes during the search, in which crossings beyond local ones were involved, making the global effect of crossings more salient.

Regarding the effect of crossing angles with set 1 drawings, subjects spent significantly more time with c3 than with c2. The eye movement data and user comments had made it clear that the longer time with c3 was caused by the slower eye movements and the extra back-forth moves at and around the crossing points, which in turn was caused by the sharp crossing angles.

In summary, how crossings affected eye movements can be summarized as follows:

1. When there were no crossings, eye movements were smooth and fast.
2. When edges crossed with large angles, eye movements were slower, but remained smooth.
3. When edges crossed with small angles, eye movements were very slow, and no longer smooth (back-forth moves at the crossing points).

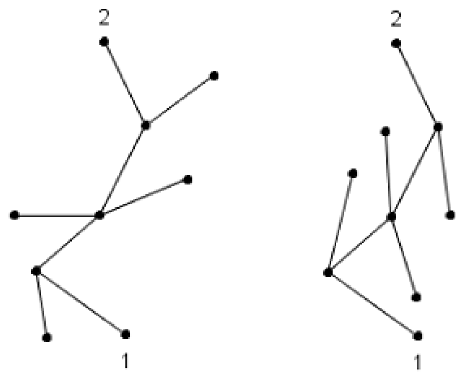


Fig. 8 Geodesic path tendency suggests that the path between nodes 1 and 2 on the left drawing is easier to detect than that on the right.

The existence of geodesic path tendency was observed in set 2 drawings. This tendency indicates that when there are branch edges pointing to the target nodes on the path that is being searched, human eyes can easily slip into those branches, causing delays in response time. The effect of this tendency can be illustrated in Figure 8. Given the drawing on the left, by changing the positions of the unlabeled nodes, the branch edges are made to go toward the labeled nodes, as shown in the right drawing. As a result of this, the path between nodes 1 and 2 becomes harder to follow.

5 General Discussion

Our research presented in this paper demonstrates the importance and feasibility of establishing aesthetics based on human graph reading behavior. In the first experiment, if we measured only task performance, we would not be able to observe that crossings actually contributed little to the time difference between crossing and non-crossing drawings. This observation has resulted in two important findings: the new aesthetic of crossing angles and the graph reading behavior of geodesic path tendency. The qualitative observations about crossing angles and the tendency have been validated in separate studies by Huang et al. [13,16], with quantitative evidence of rigorous controlled experiments.

Further, the establishment of the aesthetic of crossing angles has led to a new theoretic research area: drawing graphs with large crossing angles, which was initiated by Didimo et al. [6] and has attracted immediate attention from the research community (e.g., [1,2,8, 11,24]). Practically, it is well known that crossing minimization is computationally difficult to achieve [5], and in many real world graphs, crossings cannot be completely removed. The aesthetic of crossing angles helps to improve the situation: we may draw graphs with a few more crossings, but achieve the same level of effectiveness by maximizing crossing angles. The finding of geodesic path tendency enriches our limited knowledge of how people read node-link diagrams. The knowledge of this behavior is useful for us to design more human-centered graph visualizations: an immediate implication is that paths between important nodes or between nodes of interests should be laid out close to the geodesic path, while irrelevant branches or paths should be put further away.

This paper also demonstrates the usefulness of eye tracking in understanding the cognitive process of humans in reading graphs. From a broader perspective, since people read graphs through their eyes, the use of eye tracking should hold great promise for the evaluation of graph visualizations.

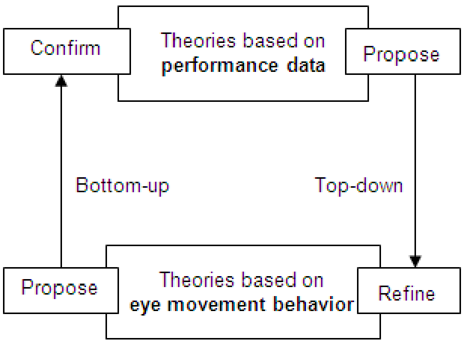


Fig. 9 Illustration of usefulness of eye tracking for the development of theories of how people read graphs.

More specifically, eye tracking can be used to obtain valuable insights beyond performance. As illustrated in Figure 9, eye tracking assists the development of human graph reading theories in both bottom-up and top-down fashions. In some cases, theories may be proposed based on performance data first. These theories can then be refined and proved based on eye movement behavior. For example, the aesthetic of crossings was first established based on performance data [29], then it was found that in our eye tracking studies that its impact can be moderated by changing the size of crossing angles. In other cases, theories may be proposed based on eye movement behavior and verified in controlled experiments based on performance data. For example, the behavior of geodesic path tendency was observed in our eye tracking studies. Then this tendency was confirmed through performance-based controlled experiments [13].

On the methodological perspective, we did not use traditional eye movement measures due to their limitations specific to our experimental purposes. Instead, eye movement videos were used, which has proved to be an effective approach in establishing aesthetics based on human graph reading behavior. Our approach can be summarized as follows:

1. Measure performance and record eye movements at the same time and analyzed them in relation to each other to answer what and how.
2. Use questionnaires and interviews while asking users to reflect and explain their eye movement strategies to answer why.

References

1. Angelini, P., Cittadini, L., Di Battista,G., Didimo, W., Frati, F., Kaufmann, M. and Symvonis, A. (2009) On the Perspectives Opened by Right Angle Crossing Drawings. In Proc. 17th International Symposium on Graph Drawing (GD'09), LNCS 5849, 21-32.
2. Argyriou, E.N., Bekos, M.A. and Symvonis, A. (2010) Maximizing the Total Resolution of Graphs. In Proc. 18th International Symposium on Graph Drawing(GD'10), LNCS 6502, 62-67.
3. Carpenter, P.A. and Shah, P. (1998) A Model of the Perceptual and Conceptual Processes in Graph Comprehension. Journal of Experimental Psychology: Applied, 4: 75-100.
4. Convertino, G., Chen, J., Yost, B., Ryu, Y. and North, C. (2003) Exploring Context Switching and Cognition in Dual-View Coordinated Visualizations. In Proc. the Conference on Coordinated and Multiple Views in Exploratory Visualization, 57-66.
5. Di Battista, G., Eades, P., Tamassia, R. and Tollis, I. (1998) Graph Drawing: Algorithms for the Visualization of Graphs, Prentice Hall.
6. Didimo, W., Eades, P. and Liotta, G. (2009) Drawing Graphs with Right Angle Crossings. In Proc. International Symposium on Algorithms and Data Structures, 206-217.
7. Duchowski, A.T. (2003) Eye Tracking Methodology: Theory and Practice. Springer-Verlag, London.
8. Dujmovic, V., Gudmundsson, J., Morin, P. and Wolle, T. (2010) Notes on Large Angle Crossing Graphs. In Proc. Computing: the Australasian Theory Symposium.
9. Duncan, J. and Humphreys, G.W. (1989) Visual Search and Stimulus Similarity. Psychological Review, 96(3): 433-458.
10. Goldberg, J.H. and Kotval, X.P. (1999) Computer Interface Evaluation Using Eye Movements: Methods and Constructs. International Journal of Industrial Ergonomics, 24, 631-645.
11. Huang, W. and Huang, M. (2010) Exploring the relative importance of crossing number and crossing angle. In Proc. the 2rd International Symposium on Visual Information Communication (VINCI'10), Article 10, 8 pages.

12. Huang, W., Huang, M. and Lin, C.C. (2011) Aesthetic of angular resolution for node-link diagrams: Validation and algorithm. In Proc. 2011 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'11), Pittsburgh, PA, USA, 18-22 September 2011.
13. Huang, W., Eades, P. and Hong, S.-H (2009) A Graph Reading Behaviour: Geodesic Path Tendency. In Proc. the IEEE Pacific Visualization Symposium (PacificVis'09). IEEE Press, pp.137-144.
14. Huang, W., Hong, S.-H. and Eades, P. (2005) Layout Effects on Sociogram Perception. In Proc. 13th International Symposium on Graph Drawing (GD'05), Lecture Notes in Computer Science, Volume 3843, Springer, 262-273.
15. Huang, W., Hong, S.-H. and Eades, P. (2007) Effects of Sociogram Drawing Conventions and Edge Crossings in Social Network Visualization. Journal of Graph Algorithms and Applications (JGAA), 11 (2): 397-429.
16. Huang, W., Hong, S.-H. and Eades, P. (2008) Effects of Crossing Angles. In Proc. the IEEE Pacific Visualization Symposium 2008 (PacificVis'08), 41-46.
17. Huang, W. and Eades, P. (2005) How people read graphs. In Proc. the 2005 Asia-Pacific symposium on Information visualisation (APVIS'05), 51-58.
18. Korner, C. (2004) Sequential Processing in Comprehension of Hierarchical Graphs. Applied Cognitive Psychology, 18(4): 467-480.
19. Korner, C. (2011) Eye movements reveal distinct search and reasoning processes in comprehension of complex graphs. Applied Cognitive Psychology. DOI: 10.1002/acp.1766.
20. Korner, C. and Albert, D. (2002) Speed of Comprehension of Visualized Ordered Sets. Journal of Experimental Psychology: Applied, 8: 57-71.
21. Krackhardt, D. (1996) Social Networks and Liability of Newness for Managers. Trends in Organizational Behavior, 3: 159-173.
22. Lohse, G.L. (1997) The Role of Working Memory on Graphical Information Processing. Behavior & Information Technology, 16(6): 297-308.
23. McGrath, C., Blythe, J. and Krackhardt, D. (1997) The Effect of Spatial Arrangement on Judgements and Errors in Interpreting Graphs. Social Networks, 19(3): 223-242.
24. Nguyen, Q., Eades, P., Hong, S.-H. and Huang, W. (2010) Large Crossing Angles in Circular Layouts. In Proc. the 18th International Symposium on Graph Drawing (GD'10), 21-24 September, Konstanz, Germany, 397-399.
25. Peebles, D. and Cheng, P.C.H. (2003) Modelling the Effect of Task and Graphical Representation on Response Latency in a Graph Reading Task. Human Factors, 45(1), 28-45.
26. Pohl, M., Schmitt, M. and Diehl, S. (2009) Comparing the Readability of Graph Layouts using Eyetracking and Task-oriented Analysis. In Proc. the Fifth International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging, 49-56.
27. Purchase, H. (1997) Which Aesthetic Has the Greatest Effect on Human Understanding? In Proc. the 5th international Symposium on Graph Drawing (GD'97), LNCS 1353, Springer-Verlag, 248-261.
28. Purchase, H., Pilcher, C. and Plimmer, B. (2011) Graph Drawing Aesthetics - Created by Users not Algorithms. IEEE Transactions on Visualization and Computer Graphics, IEEE Computer Society.
29. Purchase, H., Cohen, R.F. and James, M. (1996) Validating Graph Drawing Aesthetics. In Proc. the Symposium on Graph Drawing (GD'95), LNCS 1027, Springer-Verlag, 435-446.
30. Ratwani, R.M., Trafton, J.G. and Boehm-Davis, D.A. (2003) Thinking Graphically: Extracting Local and Global Information. In Proc. the Twenty-Fifth Annual Conference of the Cognitive Science Society.
31. Ratwani, R. M., Trafton, J. G. and Boehm-Davis, D.A. (2008) Thinking Graphically: Connecting Vision and Cognition during Graph Comprehension. Journal of Experimental Psychology: Applied. 14(1), 36-49.
32. Rayner, K. (1998) Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124(3): 372-422.
33. Ware, C., Purchase, H., Colpoys, L. and McGill, M. (2002) Cognitive Measurements of Graph Aesthetics. Information Visualization, 1(2): 103-110.
34. Wasserman, S. and Faust, K. (1994) Social Network Analysis: Methods and Applications. Cambridge University Press.