On the Effects of Visual Cues in Comprehending Distortions

A. Zanella, M. Rounding, M.S.T. Carpendale
Department. of Computer Science, University of Calgary,
2500 University Dr NW, Calgary, AB.,
T2N 1N4
{azanella,rounding,sheelagh}@cpsc.ucalgary.ca

ABSTRACT

As the primary metaphor for computer use shifts from an extension of the personal desktop to a gateway into a vast information space, representing this expanse of information on our relatively small screens is becoming increasingly problematic. One possible solution for this screen real estate problem is to make multi-scale presentations by magnifying areas of interest and compressing others. The creation of these presentations makes use of some form of distortion. Distortion in turn changes the way in which information can accurately be read. In this paper we describe a study about relative difficulty in reading distortions. We investigate the effect of introducing viewing cues such as the cartographic grid and shading on people's ability to interpret distortions. We look at two interpretation issues: whether people can locate the region of magnification and whether people can read changes in degree of magnification of these regions. We present the findings of this study and a discussion of its results.

Keywords

User study, distortion viewing, viewing cues, information visualization, detail-in-context magnification

INTRODUCTION

All too often, when viewing information on a computer, it is the size of the screen on which the information is displayed that is the limiting factor. This can be true whether one is viewing a single image or map, coping with multiple files when editing or coding, or trying to organize the windows and icons that are necessary for one's current task. In fact, computational advances over the last twenty-five years have intensified this problem. Processing power and storage capacity have increased in leaps and bounds. In comparison, the sizes of our display screens have inched outwards. This discrepancy between a computer's display space and its information space has been called the screen real estate problem and is associated with problems in navigation, interpretation and recognition of relationships between items [7].

One recurring theme in screen real estate research has been the use of various types of distortion. In some regards this is not surprising since traditionally 'distortions' have often been used to address problems of fitting a set of information into a given space and to provide the desired information emphasis. For example, carefully chosen 'distortions' or projections are used in the creation of flat maps from our spherical globe and many illustrations and diagrams carefully present selected regions of information subtly enlarged to best elucidate the chosen message.

There has been considerable discussion around the advantages and disadvantages of presentation methods that make use of distortion on the computer screen. There are studies that attribute advantages such as reduced cognitive load [1] and improved navigation [6, 14] to distortion based methods. However, there has not been widespread acceptance. This may be due to factors such as general discomfort with the use of distortion and perhaps the lack of awareness of its widespread use in other presentation mediums. More likely it is due to the fact that on a computer we, as users, are faced with interactive distortions and therefore we have the task of establishing just the right use of distortion to create just the right information emphasis. Perhaps even more crucial is that fact that we, as users, are also faced with the interpretation or perception issues involved in reading these distorted presentations.

To better understand these issues, we ran a controlled user study with the goal of gaining a better understanding of how to support people's interpretation of distortions. We examine the effects of four different levels of viewing support: no cues, the cartographic grid, shading and the use of shading and the grid in conjunction. In the next section, we describe some of the underlying ideas

A. Zanella, M. Rounding, M. S. T. Carpendale, On the Effects of Visual Cues in Comprehending Distortions. *Technical. Report No. CPSC-2000-668-20, Dept. of Computing Science, University of Calgary*, Sept., 2000

concerning the comprehension of distortions. This is followed by a description of our user study. We then conclude with a discussion of the results, and pointers to future work.

RELATED WORK

Research into more effective use of current displays has been categorized as either distortion based or non-distortion based [7]. Non-distortion based screen real estate research has led to most of the more frequently used computational presentation paradigms such as windows with pan, scroll and zoom. However, no one claims, at least in their current manifestations, that this is the perfect solution. Many researchers have noted limitations of access through pan, scroll and zoom, such as getting lost in information spaces [8], problems with maintaining context when examining information details and interpretation issues in comparisons across disparate information spaces [3, 15].

Ideally one would like to be able to take advantage of our natural visual pattern recognition abilities by being able to see the entire image. However, it is also important to see areas of interest in sufficient detail, and to be able to relate these details to their immediate surrounds and to their global context. This desire has fueled considerable research, pioneered by Spence and Apperley's Bifocal Display [15] and Furnas' [4] paper on Generalized Fisheyes. Subsequently several presentation methods have been developed [2, 5, 13] that create displays that vary considerably visually and algorithmically (for surveys see [7, 10]). Research towards the development of detail-in-context methods has concentrated on visual capabilities, for instance, the number and type of foci and the degree of magnification.

These techniques are said to support human potential for visual gestalt, to reduce cognitive effort needed for the re-integration of information across separate views and to address navigational problems by accessing spatial reasoning. Also studies indicate that setting detail in its context is common practice in human memory patterns [5] and that there is increased user performance in path finding tasks [6, 14]. However, there has not been widespread acceptance. This may be due to the fact that all of these methods make use of some form of distortion. Though visual communication issues have motivated research in this area, new comprehension issues continue to arise. These include problems recognizing that the information has not changed in creating the distorted presentation, problems interpreting the information in its distorted form, and general discomfort with the use of distortion.

While comprehension problems led to the creation of detail-in-context presentations new comprehension issues have arisen within the research [3, 9]. There have been comments about users having trouble recognizing that the altered presentations held the same information [9], about user disorientation [6, 12] and about how distortions may interfere with the users mental map [9, 16]. It has been suggested that limiting the distortions to preserve orthogonality, proximity and topology may help to support a user's mental map [9]. Subsequently it has been noted that a choice may have to be made between distortions that preserve proximity and distortions that preserve orthogonality [16]. Van der Heyden et al. [17] continue his line of reasoning.

EMPIRICAL STUDY

Overview

We are interested in understanding how to better support the interpretation of distortions. As just discussed there are many types of distortions and many types of information representations that might be usefully viewed through a distortion presentation. For this study we limited the distortions to three sizes of detail-in-context lenses. A detail-in-context lens provides space for a region of increased magnification by compressing the immediate surrounds. These lenses integrate the magnified detail with their context providing a detail-in-context presentation. The left hand-side of Figure 1 shows a single lens and the right hand side shows a group of three lenses of the three sizes used in the study. These lenses are all constrained in that the distortion does not spread across all of the context.

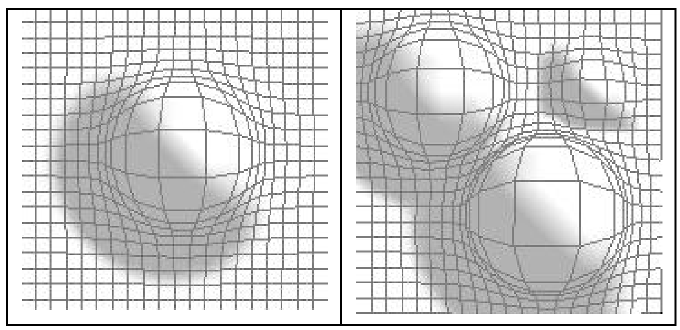

Figure 1: detail-in-context lenses

Because of the possibility that the nature of the information representation may well affect people's ability to interpret the lenses we chose to avoid either limiting the study to a particular representation or including too many variations. We chose to use maps, since they are a relatively familiar type of representation. Within this limitation we chose six maps that represent a common range of map types (Figure 2).

In studying how to provide sufficient support to make changes in scale visually explicit, we examine the use of *visual cues*. The term visual cue is used to indicate any aspect of the display that has been added for human perceptual reasons, such as attracting attention, creating emphasis, or adding explanation, rather than to directly represent some aspect of the information. In our study we

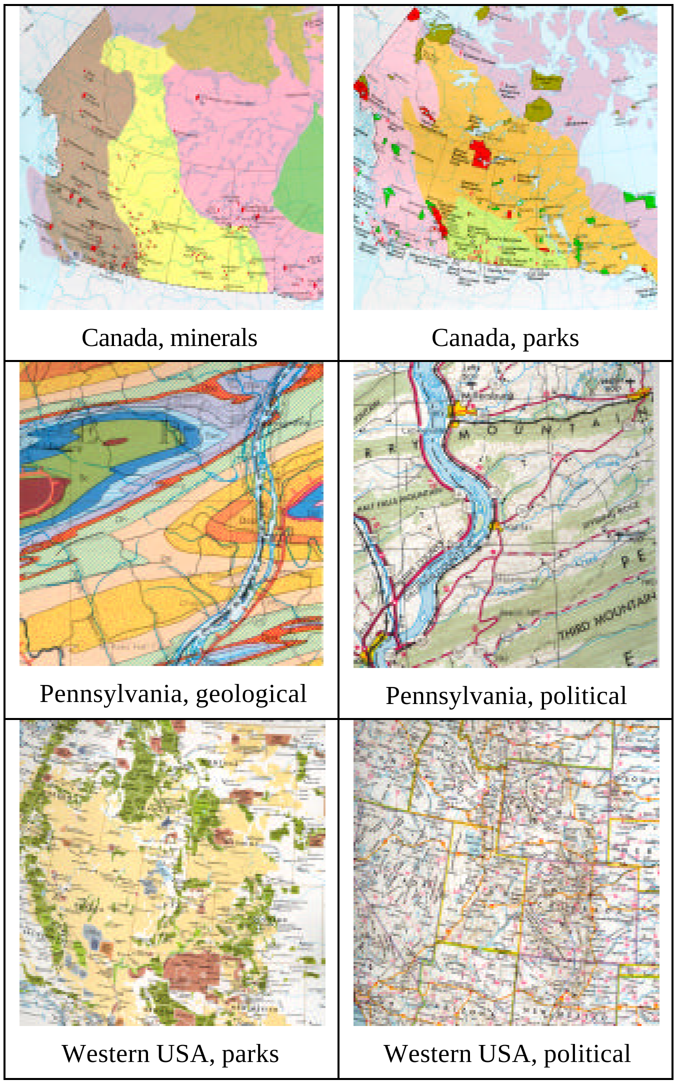

Figure 2: the maps, with no lenses

have four conditions of visual cues: the grid, shading, the grid and shading combined and no visual cues. Figure 3 illustrates these four conditions.

We created three pairs of images from each map so that there would be a pair for each map with each of one, two and three lenses. In each pair we included one image without visual cues and another with one of the other three possible cue combinations. The result of this was that each participant saw one map with the same set of lenses twice: once with visual cues and once without. Which one they saw first was not controlled and was determined randomly by the software at test-time.

We recruited thirty participants, which were mostly computer science students in undergraduate and graduate levels. Other participants included graduate students in other areas. They were half males and half females.

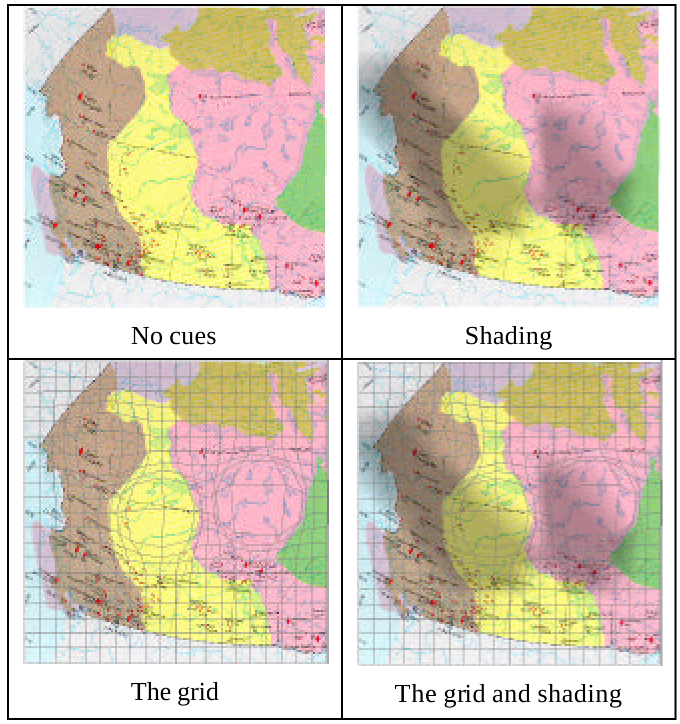

Figure 3: The four visual cue conditions applied to the same map (Canada, minerals) with the same three lenses in each image

Experimental Task

Thirty-six images were presented to each participant, each containing from one to three lenses. Each lens configuration on a given map was shown twice; once with no cues and once with one of the three visual cues. For each participant the entire sequence of images was randomized. Participants were asked to locate and rank the lenses simultaneously. They indicated the location of the lens by clicking in the centre of the area of maximum magnification. They indicated their impression of relative degree of magnification by the order in which they clicked on the lenses. For example the first lens they selected was considered to be their choice for the largest lens and subsequently the rest of the lenses were to be selected in order of decreasing size.

Procedure and Data Collection

Participants were greeted, given a brief overview of our research and filled out a background questionnaire that assessed their experience with computers and digital images. Based on the answers to this initial questionnaire the participants were divided into two groups: one with participants that had little or no experience with the lens system, and another with participants that had seen the lenses before. We emphasized the difference in experience by giving the second group a short training session. They were given approximately twenty minutes to familiarize themselves in the idea of lenses in an interactive setting. The trained participants were shown how to create lenses on maps and how to interact with them. This interaction included a demonstration of the different visual cues that could be used to accentuate the graphical lenses as well as changing lens magnification.

For the training sessions, participants got to use the same software we used to create the images. This software allows loading different images and applying different lenses to them. The lenses can then be moved around. Explicit control of the visual cues is also available, allowing the participants to play with lenses while removing and adding the different cues. We also showed the trainees how to change the magnification of a lens and guided them through several simple tasks where lenses were created and manipulated.

The goal was to have the participants who had received the benefit of this short training session leave with a clear understanding of how the lenses and how visual cues behaved in an interactive sense. We were interested in whether the participants could more immediately identify the graphical lenses if given a more extensive introduction to the concept.

All the participants received a careful explanation of the experimental task with the aid of a visual tutorial. Then they were asked to find and rank the lenses in the images in descending order of magnification. We designed the software to track participants' progress through the image set. Each time a user clicked on a particular image, the program stored the (x, y) coordinates in pixels that the participant clicked on, plus the distance to the center of the nearest lens. It also tracked the cumulative time the participant took in finding the lenses in the image. This was accomplished by recording the time between when the image appeared onscreen and when a click was recorded.

At the end, we asked participants perform the same experimental task with a small set of images using a talk aloud procedure. Here they performed the same task but explained to the experimenter their process and strategies. These sessions were video taped.

Participants filled out a post-session questionnaire where they indicated their impressions and preferences in terms of locating and ranking the lenses. Additional space was provided for comments.

The independent variables are: the visual cue condition (no cue, grid, shading, grid and shading, see Figure 3); and the information representation (Canada-minerals, Canada-parks, Pennsylvania-geological, Pennsylvania-political, Western US-parks, Western US-political, see Figure 2).

The dependent variables are the location information, the rank order, and the time used. For location information we kept track of the (x, y) point that was actually selected, the location of the closest lens, and the distance to the closest lens. Further, we also recorded whether the point selected was in the focal region, in the periphery of the focal region, within the lens or elsewhere in the image but just closer to this lens than any other lens. Relative rank information for the lenses within an image was also tracked. There is no rank information for images with only a single lens. Also, the time taken over each lens was recorded.

Qualitative data was also collected; we asked users about their preferences for the visual cues when locating and ranking the lenses and the sessions were video taped.

RESULTS AND DISCUSSION

Locating Lenses and Visual Cues

For the task of locating the lenses, all the visual cues improved the situation significantly. Figures 4, 5 and 6 compare each visual cue, grid, shading and grid plus shading respectively with no visual cues. These charts show the distances in pixels of the points clicked from the closest lens' centre. Distances of within a radius of twenty pixels are within the lens focal region. Distances between twenty and forty are on the periphery of the focal region. Distances of less than sixty can be said to be within the lens. That is they are somewhere within the region that is affected by the distortion. We capped the entirely missed lenses and extremely large distances. Essentially the flat line of the capped region represents the number of lenses that were not found. Each chart compares the differences in visual cues only, with the map and particular lens combinations held constant. Figure 7 compares the three visual cues. In comparison with each other the grid is significantly better than either shading or the grid and shading combined.

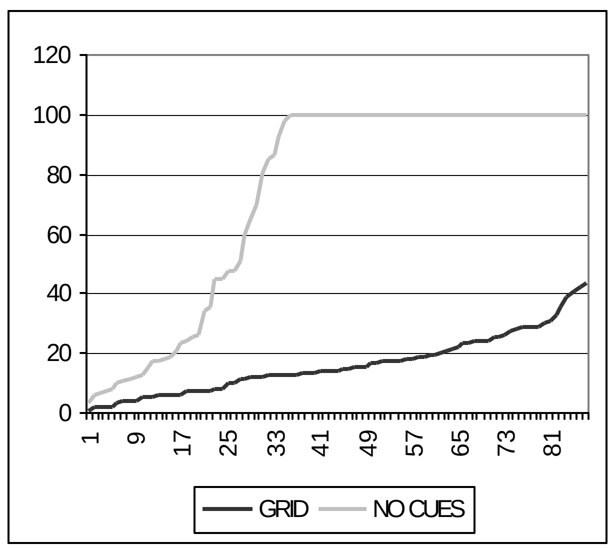

Figure 4: Contrasting location accuracy, the grid versus no cues

The Maps: differing representations

Note the differences in distance curves in Figures 4, 5 and 6 for the no cue condition. This is due the change in map type. Figure 8 compares the location accuracy across the six maps. To be able to look at the differences between maps the results shown in Figure 8 are all from the no cue condition. There are both some striking similarities and some surprising differences. As might be excepted both the political maps, Pennsylvania-political and Western US-political seem to pose the same level of difficulty. These political maps have a considerable amount of text and roads and boundaries. Similarly both the categorical maps, Canada-minerals and Pennsylvania-geological, produce comparable location results. In contrast, the results for the two parks maps are significantly different from each other and from the other maps. People found of these maps very different. With the Canada-parks map, by far the most lenses are entirely missed, while with the Western US parks map the least number of lenses are actually missed.

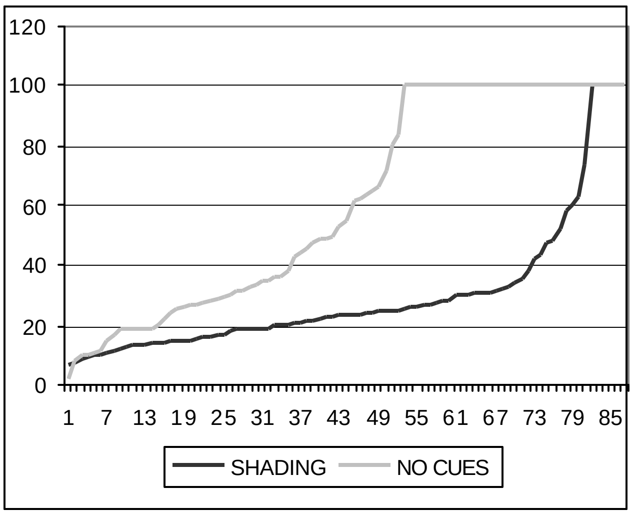

Figure 5: Contrasting location accuracy, shading versus no cues

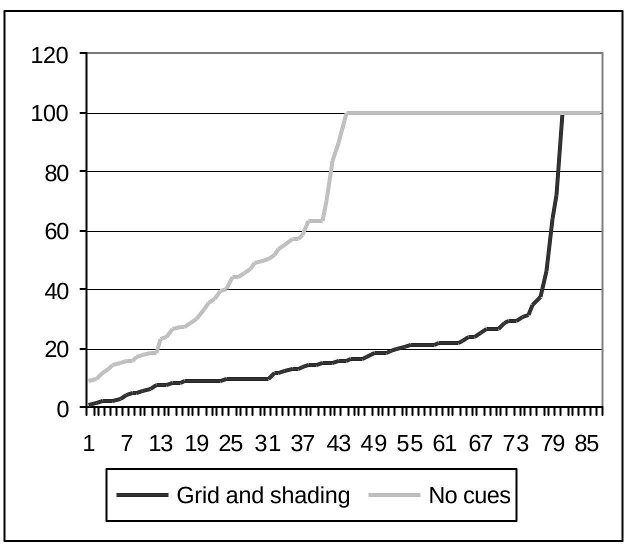

Figure 6: Contrasting location accuracy, the grid plus shading versus no cues

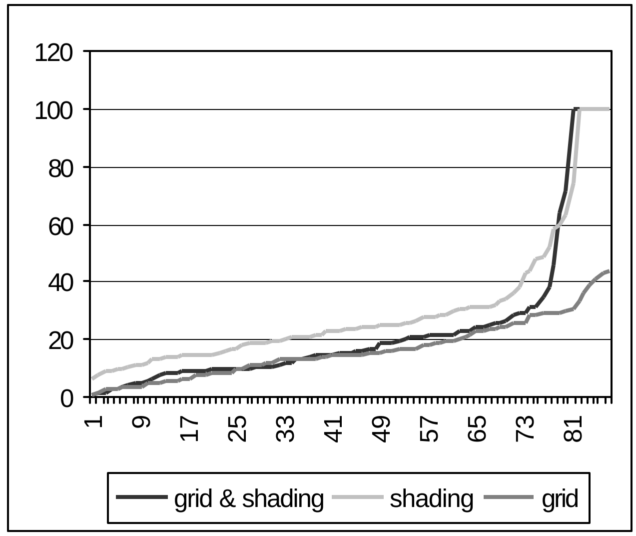

Figure 7: Comparing the three visual cues grid, shading and grid plus shading for location accuracy

Ranking Lenses and Visual Cues

While the results for the locating task are much as expected the results for ordering the lenses according to degree of magnification are not. Ranking errors were counted by lens. For instance, suppose the lenses were ordered so that the smallest was indicated to be the largest and the largest was indicated to be the smallest (large, middle, small ordered as small, middle, large). This would be counted as two ranking errors since the middle one is where it should be. No ranking errors were possible if there was only one lens in the image. Other errors that were also counted are missed lenses and extra lenses. Extra lenses are those that were indicated by a participant when the closest lens has already been selected. Figure 9 shows these errors as percentages of the possible errors.

With the no cue condition the results were not surprising. Approximately a quarter of the lenses are not noticed and of those noticed approximately a quarter are mis-ranked. With the grid or the grid plus shading very few lenses are missed and considerably fewer are mis-ranked. The surprising result is that shading not only provides no support for the judging of relative magnification, it appears to be very misleading. The overall error count for ranking errors with shading is close to fifty percent. However, with a couple of the representations, notably Pennsylvania-geological (Figure 10), virtually all participants had some ranking errors.

Learning Effects

There was no significant learning effect resulting from our training sessions in any of the conditions except for shading. The participants who had received a training session made use of the shading cue much more accurately (p=0.001).

As far as the other cues are concerned it is possible is that our training session was not adequate to provide a significant learning effect. For example, we did not explain that in a distorted image with a grid they had to look for the lines bending to be able to find the center of the lens.

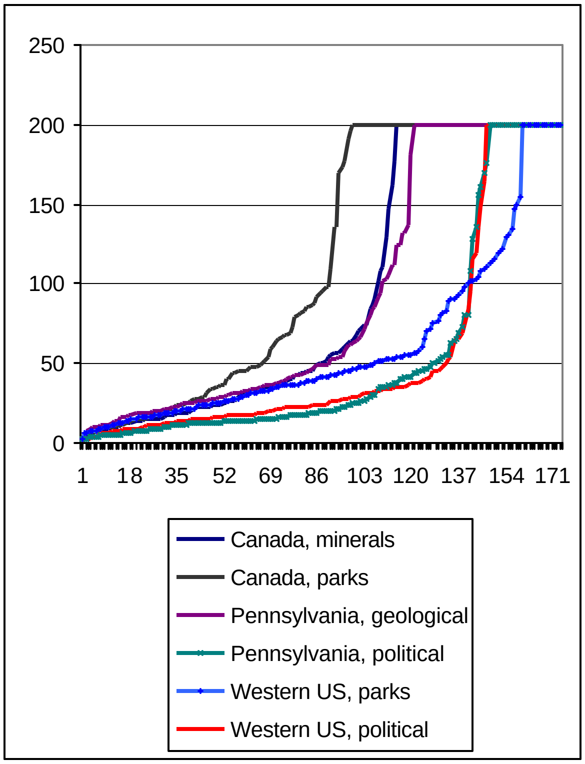

Figure 8: Comparing the maps for location accuracy

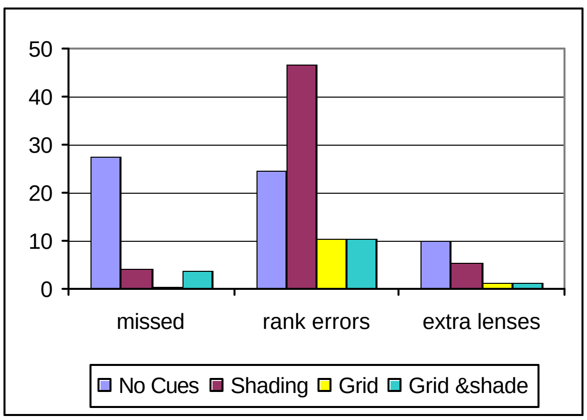

Figure 9: Contrasting the different error types under the different visual cue conditions

Participants Styles and Preferences

During a debriefing session, we asked the users to go through a set of six to ten more images while relating out loud their thoughts and actions while finding and ranking the lenses. Sometimes we prompted the users for more information with questions such as: "Why do you think it is easy to find the lens in this image?" or, "What type of cue do you first look for in an image?" In this debriefing we wanted to capture the participants' strategies as a more qualitative type of data as compared to *what* they did as measured by the software.

The participants' preferences were recorded in the post-session questionnaire. Table 1 illustrates the results of this questionnaire. The grid by itself was most preferred by our group of participants for the purposes of both locating and ranking lenses.

	Locating	Ranking
No Cues	0	1
Grid	16	20
Shading	1	1
Grid and Shading	13	8

Table 1 – Preferences in locating lenses

For the purpose of locating lenses, the grid cue was preferred for sixteen participants, followed by grid and shading which was named as the preferred cue for thirteen participants. Only a single participant pointed to shading only as their preferred cue for finding lenses.

For ranking, the grid was clearly the preferred cue with twenty participants choosing it over only eight participants who chose grid and shading and the single participant who chose shading. Interestingly, there was also one participant who chose no cues in this category. The participant justified this by saying that they were looking for distortions of word sizes to help them ascertain the different magnifications.

CONCLUSIONS

All of the visual cues significantly improved people's ability to locate lenses. By far the most useful and most preferred cue was the grid. Even when combined with shading, the grid alone still allowed people to perform better overall. We suspect that the reason the grid performed so well overall was that the grid lines were actually distorted as well. This makes the information easier to read.

It was interesting to find out that the shading cue did not actually provide that much useful information for our participants.

According to cognitive science literature the ability to read shape from shading is a pre-attentive ability. "The human visual system is capable of quickly and accurately establishing three dimensions from variations in luminance [shading] only"[14]. It is probable that this ability to recognize shape from shading is one of the most primitive abilities. Ware [20] points out that distinguishing shape from shading is part of what he terms a sensory language, that bridges cultures and does not have to be learned. One would think that such a low-level visual routine would be perfect for our purposes and provide us with a method to make distortions explicit. Making use of shape from shading should ensure that it is pre-attentive abilities that are being accessed instead of possibly increasing cognitive load.

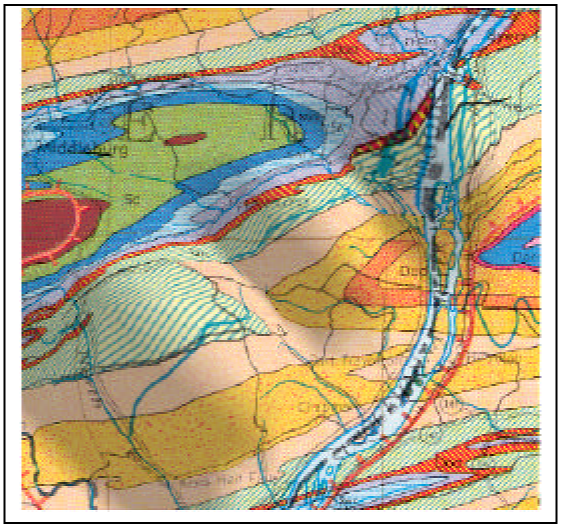

Figure 10: Pennsylvania-geological map with three lenses, a particularly difficult one for ranking lenses

In contrast to all this in our study the effects of shading as a visual cue are counter intuitive.

- Shading did improve participant's abilities to locate the lenses. This is in keeping with the expectations from the cognitive science literature mentioned above.
- Shading however was the only cue that showed a significant learning effect. This is in stark contrast to the expectations that shading will behave as a pre-attentive ability. If people are learning to recognize the effects of shading then it is not accessing low-level routines and is probably being cognitively processed.
- Shading went beyond not helping with participant's efforts to rank the lenses in order of decreasing magnification. It definitely seemed to confused people as to relative size.

For results as surprising as this one needs to consider possible causes. Some of the suspicions we have are:

- Gouraud shading may not be an adequate substitute for actual shading. If people are not seeing the addition as shading it may simply have the effect of darkening the colours.
- Looking at a 3D image on a flat screen is not the same as viewing 3D objects in the real world. We need to be careful about assuming that all abilities will transfer.

This study has demonstrated the effectiveness of using visual cues to aid in comprehension of distortions. It has also raised several questions about the robustness of the use of shading on a computer screen.

ACKNOWLEDGMENTS

We thank D. Cowperthwaite and M. Tigges for the software used to create the images and S. Greenberg for helpful suggestions throughout the project. We would also like to acknowledge all the participants in the study as well as the friends and colleagues who volunteered their time to help us develop our study.

REFERENCES

1. Bartram, L., & Ho A., & Dill, J., & Henigman F. (1995). The continuous zoom: A constrained fisheye technique for viewing and navigating large information spaces. In *Proceedings of the ACM Symposium on User Interface Software and Technology: UIST'95*, pp. 207-215, Pittsburgh, USA.
2. Carpendale, M. S. T., Cowperthwaite, D. J. and Fracchia, F. D. Three-Dimensional Pliable Surfaces: For effective presentation of visual information, *In UIST'95: Proceedings of the ACM Symposium on User Interface Software and Technology*, (Pittsburgh USA, Nov. 1995), ACM Press, 217-226.
3. Carpendale, M. S. T., Cowperthwaite, D. J. and Fracchia, F. D. Making Distortions Comprehensible, *In IEEE Symposium on Visual Languages*, (Sept. 1997), IEEE Computer Society Press, 36-45.
4. Furnas, G. W. (1986). Generalized fisheye views. In Mantei and Orbeton, ed., *Proceedings of the ACM Conference on Human Factors in Computer Systems: CHI'86*, pp. 16-23.
5. Hamel, J. &, Michel, R., & Strothotte, T. (1996). Visibility through inaccuracy: Geometric distortions to cluttering in route maps. In *Proceedings of the Central European Conference on Computer Graphics*, pp. 123-132.
6. Hollands, J. C. & Carey, T. T., & McCann, C. A. (1989) Presenting a graphical network: A comparison of performance using fisheye and scrolling views. *In Designing and Using Human-Computer Interfaces and Knowledge Based Systems*, pp. 313--320.
7. Leung, Y. K., & Apperley, M. D. (1994). A review and taxonomy of distortion-oriented presentation techniques. ACM Transactions on Computer-Human Interaction, vol. 1(2) pp.126-160.

8. Meister, D. (1989). Lost in computer space. International Journal of Human-Computer Interaction, vol.1(1) pp. 5-21.
9. Misue, K., & Eades, P., & Lai, W. & Sugiyama, K. (1995). Layout adjustment and the mental map. *Journal of Visual Languages and Computing*, vol. 6(2) pp.183-210.
10. Noik, E. G. (1994). Space of Presentation Emphasis Techniques for Visualizing Graphs, *In the Proceedings of Graphics Interface'94*. pp. 225-233.
11. Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, vol. 331(14), pp.163-166.
12. Sarkar, M. & Brown, M. H. (1994). Graphical fisheye views. *Communications of the ACM*, vol. 37(12), pp. 73-84. ACM Press.
13. Sarkar, M. & Snibbe, S., & Tversky, O. J., & Reiss, S. P. (1993). Stretching the rubber sheet: A metaphor for viewing large layouts on small screens. In *Proceedings of the ACM Symposium on User Interface Software and Technology: UIST'93*, pp. 81—91. Atlanta, Georgia, USA, ACM Press.
14. Doug Schaffer, D., & Zuo, Z., & Bartram, L., & Dill, J., & Dubs, S., & Greenberg, S. & Roseman, M. (1993). Comparing fisheye and full-zoom techniques for navigation of hierarchically clustered networks. In *Proceedings of Graphics Interface*, (1993), 87-96.
15. Spence, R., & Apperly, M. Data base navigation: an office enviroment for the professional. *Behaviour and Information Technology*, vol. 1(1), (1994). 43-54.
16. Storey, M.-A. D., & Muller, H. A. Graph layout adjustment strategies. *In the Proceedings of the Third Symposium on Graph Drawing, vol. 1027 of Lecture Notes in Computer Science*, (1995), 487-499.
17. J. van der Heyden, K. Inkpen, S. Atkins and M. S. T. Carpendale. A User Centered Task Analysis of Interface Requirements for MRI Viewing. *In Graphics Interface*, (June, 1999) 18-26.
18. Ware, C. The foundations of experimental semiotics: a theory of sensory and conventional representation. *Journal of Visual Languages and Computing*, vol. 4, (1993) 91-100.