Perceptual Invariance of Nonlinear Focus+Context Transformations

Keith Lau
Department of Computer Science
University of British Columbia

klau@cs.ubc.ca

Abstract

Focus+Context techniques are commonly used in visualization systems to provide both the details and the context. The "shaker paradigm", in which an image is rapidly alternated with a transformed image, was recently developed to find the threshold of perceptual invariance on scaling and rotations. Using this framework, we aim to examine the limits of perceptual invariance on fisheye transformations. We find that there is a "no-cost" region where visual search for a given target is unaffected by an abrupt, noticeable fisheye transformation. There is also a "low-cost" region where transformations of larger magnitude result in performance degradation. We also find that a rectangular grid background neither aids nor impairs performance. Our findings quantify the perceptual cost of nonlinear transformations, confirming that there exist cases where increased context can be experienced with minimal or no cost. These findings contradict the belief that sudden changes are necessarily detrimental to performance, suggesting that smoothly animated transformations between visual states are not always necessary.

1   Introduction

When the viewable area of the screen is held constant, it seems inevitable that as one focuses on the specifics, the mental model of the overall picture diminishes simultaneously. This trade-off has been problematic in visualization systems, and especially so when large datasets are involved. One of the goals in the field of information visualization is to provide just enough context, so that a mental model of the whole dataset can be maintained throughout.

Many systems use "Focus+Context" techniques, in which a nonlinear transformation is applied to the image to provide both the focus and the context in one integrated image [Furnas 1986; Leung and Apperly 1994; Rao 1997; Munzner 2003]. Some techniques are able to make clear the region of focus, at the expense of intense distortion in the neighbouring regions [Sarkar and Brown 1994]. Despite these kinds of associated distortions, Focus+Context techniques are widely accepted as beneficial [Kobsa 2003; Risden 2000]. However, their costs and the effects on visual perception have not yet been quantified. In particular, we aim to determine how much visual processing is disturbed by sudden nonlinear transformations. Another goal is to find out whether adding visual cues like grid lines to the background would have an effect on the task.

1.1   Focus+ Context Motivation

Ware classifies visualization systems into three domains: (i) spatial scale, as found in geographic or biological visualization systems, (ii) temporal scale, like systems that visualize traffic patterns, and (iii) structural scale, where the user navigates in an abstract structure, such as the code of a complex computer system [Ware 2000 (Chap. 10)]. The common issue that persists among all three variations is the loss of context as the user focuses on a particular portion of the dataset.

Assuming that the user is using a small window to "move around" the dataset, the severity of the problem elevates as the size of the dataset increases. Exploring a large dataset through a small window certainly does not give much explicit information about anything else other than what is currently on the screen. Furthermore, keeping track of the navigation history is a highly demanding task cognitively [Zhang 1991]. Unfortunately, without any additional aid, people are often required to do backtracking in the attempts to remember where they have visited. An attempt to resolve the issue of a small window is to utilize multiple windows that are connected with transparent overlay [Card et al 1994]. In this case, the region of focus is in a larger size while being linked visually to its context. However, this method tends to lead to clustering very quickly.

Alternatively, in the NV3D systems, Parker et al. attempted to remedy the problem from another angle using rapid zooming [Parker et al 1998]. Instead of providing context, this method works on providing smooth transitions from the context to the object in focus. Although it only allows a part of the dataset to be visible in any given instant, the smooth animated transition was designed to aid the user by avoiding sudden jumps between views. We will show later that the theory that underlies this design may not necessarily hold.

Another way to provide a convenient mental model of the overall view is to add an overview window that always shows the current location relative to the overall view. This works by explicitly providing the context and reducing the cognitive load of the user. However, one disadvantage is that switching between the two views might be quite distracting for the mind, as the user constantly tries to relate the small overview window with the main one. Also, the issue of navigational history is still not addressed by this approach.

The solution of creating a single integrated view showing the details surrounded by context seems most preferable. If done correctly, the complete context could be preserved, without the need for artificial features like animated transitions or multiple

views. One risk, though, is having a distortion that is overly done, which can render neighbouring context unrecognizable. Nonetheless, this method of nonlinear image transformation seems highly regarded, and has been researched extensively under various names such as Focus+Context [Rao 1994], fisheye views [Furnas 1986; Sarkar 1992] distortion-oriented presentation techniques [Leung and Apperly 1994], nonlinear distortion [Keahey 1997], pliable surfaces [Carpendale 1995], and elastic presentation spaces [Carpendale 2001]. The Hyperbolic tree browser is an example that uses a variant of nonlinear transformation by magnifying the node of interest and shrinking other nodes of the tree. Some of these systems also use the metaphor of a stretchable rubber sheet [Sarkar 1993], which has been used in numerous systems. Among others, it was first employed by the Document Lens [Robertson 1993] where a rectangular lens was used for the distortion. Much more recently, the TreeJuxtaposer system used the metaphor to enable fast structural comparison of large trees [Munzner 2003].

While systems utilizing nonlinear transformations seem refined, though not perfect, many questions are open for discussion. First of all, an optimal design has not yet been specified. Second, it is not known whether any transformation is superior to others.

1.2 Transformational Invariance

The term "mental rotation", as coined by Shepard and Metzler, implied that the representation of objects is orientation-specific, which could be mentally rotated when needed [Shepard and Metzler 1971]. Later, size scaling has also been found to function similarly [Bundesen and Larsen 1975]. These discoveries consequently suggest that rotation and scaling are "natural" transformations, in that they seem to directly map to operations of the visual system.

Deriving from these theories is the belief that immediate changes in orientation or size would be hard to track. To address this issue, Markinlay et al. developed a rapid-navigation technique called *point of interest navigation*, which provides animated smooth transitions between points in the image [Markinlay et al, 1990]. However, recent findings show that this view may not be accurate. Using the "shaker paradigm", in which an original image rapidly alternated with the image transformed by rotation or scaling, interesting results pointing to transformational invariance were found. The effect of the distortion was determined by comparing the search speed for a given kind of target item in a static image against the speed found for the shaker condition. Performance was found to be dependent on the magnitude of the abrupt transformation. For rotations of 17 degrees or less, performance was unaffected; for rotations of 30 degrees or above, speed was reduced significantly. This pattern indicated at least some invariance to rotation. When experimenting with scaling, performance was unaffected for rapid size changes of up to a factor of 2. The existence of this "no-cost zone" indicated that sudden changes could indeed be tracked without any performance penalty.

Taking the shaker paradigm one step further, we will now apply it to fisheye transformations. If the paradigm could

indeed be applicable to any arbitrary transformation, it would be possible, in the future, to map out the extent to which "no-cost" regions exist for all the major kinds of transformations used in visualization systems.

2 Experiment Protocol

In this experiment, we aim to locate the "no-cost" and "low-cost" regions. Since we only have 4 levels of magnifications, we aim to first get a good approximation of the regions in this study, leaving the specific thresholds for further investigation. We also seek to determine whether a grid background has an effect on performance.

2.1 Subjects

A total of 24 subjects participated in the study. 12 subjects were shown the background grid condition, 12 were shown the images without grids. The subjects in this study were students at the University of British Columbia between ages 18 to 35, with normal or corrected-to-normal vision. During the experiment, the subjects were seated for a viewing distance from the display of approximately 55cm, and kept their hands on the keyboard with their right index fingers on the "p" key and their left index fingers on the "a" key.

2.2 Images

The images displayed to the subjects are filled with randomly-positioned items, where an item is either a T or an L (Figure 1). There are three classes of images, which contain 16, 24 and 32 items respectively. These numbers are chosen after numerous test runs to ensure that subjects will not be worn out by a task that is too difficult (when there are too many items) or a task that is too effortless (when there are too few items). The density of the items is kept consistent by adjusting the area of the image. The three image sizes used are squares that subtend visual angles of $8.5^?$, $10.5^?$, and $12.5^?$. In the grid conditions, the space between the grid lines subtends a visual angle of $0.5^?$.

Each item subtends a visual angle of $1.0^?$, given the controlled viewing distance is approximately 55 centimetres between the subject and the displayed image. Items can be in any one of the four possible orientations: $45^?$, $135^?$, $225^?$ or $315^?$. They are *locally rigid*, meaning that their shapes are not distorted by the global transformation used to place them, and their front faces are always parallel to the viewer.

In the control condition, each trial only contains one image (as shown in Figure 1a). In the experimental conditions, each trial contains a pair of images that alternate with a period of 480 ms: the static image and one of the three possible transformed images (as shown in Figure 1).

The images were generated beforehand with an OpenGL program. They were created such that the eye point was located directly above the image plane, and the height values were transformed by the magnification factor c as in the following equation:

$$z = ce^{-(x^2+y^2)}(c-c(x^2+y^2)) \qquad (1)$$

This equation is responsible for creating the fisheye effect with characteristics such that the "belching-out" shape around the center and the gentle slope near the edges, which are standard features found in all fisheye transformations [Keahey 1997; Leung and Apperly 1994]. For the three levels of transformation, call them level 1, 2 and 3 in the order of increasing magnification, level 3 is defined to be the maximum possible magnification. In other words, at level 3, increasing the magnification factor c any further would push the transformed plane beyond the eye point. In our experiment, such c value is 1.5. For the static case, c is equal to 0. Therefore, the domain of 0 to 1.5 encapsulates all points from the least to the greatest magnification level possible. Using an increment of 0.3, we have omitted the values 0.3 and 0.6 in our analysis because they were found to be very close to the static case during our test runs. Therefore, level 1 and 2 have c values equal to 0.9 and 1.2 respectively.
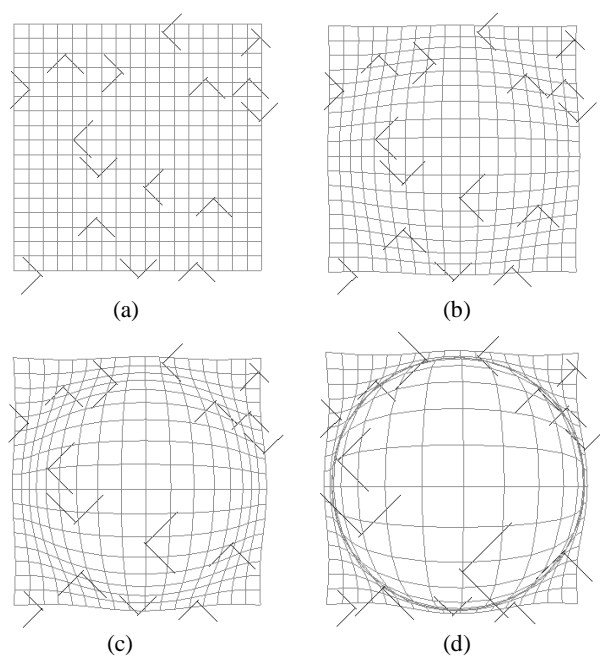

Figure 1: Images shown to users. a) Static image.
Transformed images: b) Magnification level 1, c=0.9.
c) Magnification level 2, c=1.2. d) Magnification level 3, c=1.5.

2.3 Setup

There are two sets of conditions. They are identical except that the first set of images has a rectangular grid in the background, while the second set does not. In each set, there is one control condition and three experimental conditions. Each condition consists of 120 trials, where one trial includes one image for the static condition, but includes a pair of alternating images for each experimental condition. In half of the trials, the target is absent; and it is present in the other half. The control case has no magnification, and each experimental condition differs in its level of magnification. Level 1, 2 and 3 correspond to $c = 0.9$, $c = 1.2$ and $c = 1.5$ respectively where c is the height parameter in equation (1). Each subject was shown a total of 480 images.

2.4 Procedure

Subjects were asked to find the T-shaped target amidst a group of L-shape distracter items in one of four possible images shown on a computer screen, as shown in Figure 1. They were told to hit the key labelled "p" if they believed the target was present, and the one labelled "a" if absent. They were also instructed to perform this task as quickly as possible, while keeping the percentage of error below 10%. Each subject was asked to do all four conditions, which were counterbalanced in terms of the order of presentation.

3 Analysis

There is a reciprocal relationship between search speed and error rate, so we aim to have the error rate to be approximately 10% for consistency and ease of comparison. (Figure 2 shows performance, and Figure 3 shows error rates.) We have chosen to eliminate outliers with more than 35% errors in our analysis. Also, only the reaction times in the correct target-present trials were used to calculate the search rates. The search rate was calculated by dividing the difference between the average correct reaction times in target-present trials of set sizes 32 and 16, by the difference between the set sizes. Observers with at least one search rate less than 15 ms/item in any of the conditions were removed from the analysis, since they could search an entire image without any alternation of the display, and would not experience any transformation effect. We obtained the *absolute search rates* by subtracting the average search rate of the control condition from the average search rates of its respective experimental conditions. Then we calculated the individual ratio of search rates of each observer by dividing the search rates of the experimental conditions by the search rate of the control condition. The *search rate ratios* shown in Figure 2 were averages of the search rate ratios of the 12 observers in each experimental condition.
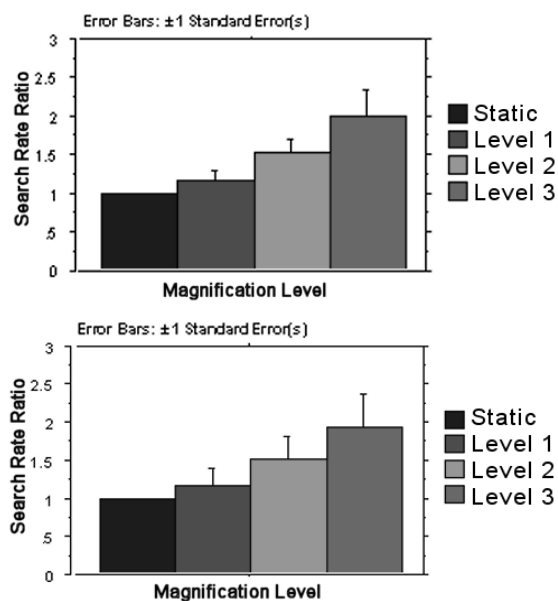

Figure 2: Search rate ratios for each magnification level.
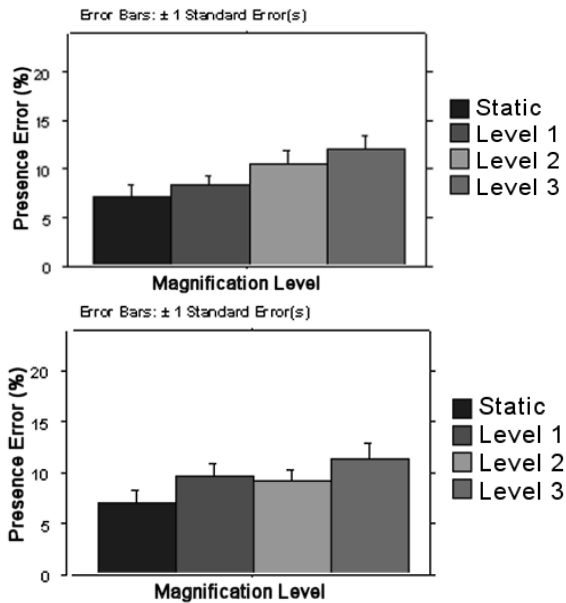Top: Background grid case. Bottom: No grid.

3

Figure 3: Error rates for each magnification level. Top: Background grid case. Bottom: No grid.

We performed a paired t-test to check if any performance degradation is significant.

Paired t-test: Hypothesized Difference = 0 (grid case)

Difference in Search Rates	p-value
between static and level 1	.752
between static and level 2	.1830
between static and level 3	.0075

We see that differences are not significant for level 1 and 2 with $p > 0.15$. Their corresponding error rates are also not significantly different (Figure 3 top); thus, we can safely rule out the issue of trade-off between error and speed. At level 3, the difference becomes significant with $p < 0.01$. Moreover, the corresponding error rate here is significantly higher. In other words, the performance degradation is apparently evident at level 3. This implies that the "no-cost" region lies around level 1 where $c=0.9$, and the "low-cost" region extends up to approximately level 2 where $c=1.2$. Similar results were obtained using data from the no-grid case.

Finally we ran another paired t-test on each search rate ratio between the grid and the no-grid case to check whether there exists a significant difference. Each case resulted in p-values $> .90$, suggesting that performance is not at all affected by the rectangular grid used.

4   Discussion and Future Work

Results show that there exists a nontrivial "no-cost" region for fisheye transformations, in which sudden transformations did not result in performance penalty. There also seems to be a "low-cost" region in which the transformation does have minimal effect on performance.

Our findings are certainly not limited to the scope of visual search tasks, since they are, in fact, representative of many common visual operations in information visualization [Ware 2000]. It would be most preferable if implications of the fisheye transformations and the shaker paradigm could be generalized to include all visualization tasks, which would be an interesting problem by itself that deserves further studies.

Another finding that is worthy of note is the zero effect of the grid on performance. A seemingly intuitive prediction might be that the grid will enhance performance in the form of a visual cue, which explicitly conveys the distorted space and reduces the cognitive load of the user. Another argument is that the grid should provide a consistent frame of reference between the static image and the transformed one, thus bridging the two. On the contrary, results show an indifference of performance. It might be so due to various reasons. Since the grid is Euclidean, while the transformation is radial, this mismatch might be enough to suppress the effect. Secondly, the grid lines might be too thin or too far apart. In any case, although there is no difference in search speed, the visual experience of the user might be quite different. This issue would require further investigation.

At a higher level, the shaker paradigm could well be used to determine which distortion operations are more natural in the mind. For those that are natural, results would show them as minimally-distracting; for those that are not, results would show drastic performance degradation.

5   Conclusions

This work has shown that the shaker paradigm [Rensink 2004] can be used to investigate the effects of nonlinear distortions in Focus+Context systems. We have proved the existence of the region where sudden transformations have no effect on the speed or accuracy of visual search. We have also established a first approximation as to where that region lies. Beyond that, performance degradation that increases with the magnitude of magnification becomes apparent. Results also show no effects of the rectangular grid on any magnitude of transformation.

6   Acknowledgements

7   References

BUNDESEN, C. AND LARSEN, A. 1975. Visual Transformation of Size. Journal of Experimental Psychology: Human Perception and Performance, 1, 214-220.

CARD, S.K., PIROLLI, P., AND MACKINLAY, J.D. 1994. The Cost-of-knowledge Characteristic Function: Display Evaluation of Direct-walk Dynamic Information Visualizations (CHI 94), 238-244.

CARPENDALE, M.S.T., COWPERTHWAITE, D.J. AND FRACCHIA, F.D. 1995. 3-Dimensional Pliable Surfaces: for the Effective Presentation of Visual Information. In Proc. ACM Symp. User Interface Software and Technology (UIST), 217-226.

CARPENDALE, S. AND MONTAGNESE, C. 2001. A Framework for Unifying Presentation Space, In Proc. User Interface Software and Technology (UIST 2001), 61-70. 10-18.
FURNAS, G., 1986. Generalized Fisheye Views. In Proc. Human Factors in Computing Systems (CHI 86), 18-23.

GREENBERG, S., GUTWIN, C., AND COCKBURN, A. 1996. Using Distortion-Oriented Displays to Support Workspace Awareness. In Proc. HCI'96 Conference on People and Computers XI, 299-314.

KEAHEY, T.A. AND ROBERTSON, E.L. 1997. Nonlinear Magnification Fields. In Proc. IEEE Symposium on Information Visualization (InfoVis), 38-45.

KOBSA, A. 2003. User Experiments with Tree Visualization Systems Technical Report UCI-ISR-03-11, University of California Irvine Department of Computer Science.

KUBOVY, M. 1986. The Robustness of Perspective. In The Psychology of Perspective and Renaissance Art, Cambridge University Press.

LAMPING, J., RAO, R., AND PIROLLI, P. 1995. A Focus+Context Technique Based on Hyperbolic Geometry for Visualizing Large Hierarchies. In Proc. Human Factors in Computing Systems (CHI 95), ACM Press, 401-408.

LEUNG, Y. AND APPERLEY, M.A. 1994. A Review and Taxonomy of Distortion Oriented Presentation Techniques. ACM Transactions on Computer Human Interaction, 1(2):126--160.

MUNZNER, T., GUIMBRETIERE, F., TASIRAN, S., ZHANG, L. AND ZHOU, Y. 2003. TreeJuxtaposer: Scalable Tree Comparison using Focus+Context with Guaranteed Visibility. In SIGGRAPH 2003 Proceedings, published as special issue of Transactions on Graphics 2003, 453-462.

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

RAO, R. AND CARD, S.K. 1994. The Table Lens: Merging Graphical and Symbolic Representations in an Interactive Focus+Context Visualization for Tabular Information. In Proc. Human Factors in Computing Systems (CHI '94), 318-322.

RENSINK, R.A. 2004. The Invariance of Visual Search to Geometric Transformation. Journal of Vision, 4. [Vision Sciences Society, Sarasota, FL, USA. May 2004.]

RISDEN, K., CZERWINSKI, M., MUNZNER, T., AND COOK, D.B. 2000. An Initial Examination of Ease of Use for 2D and 3D Information Visualizations of Web Content. Int. J. Hum.-Comput. Stud. 53(5): 695-714.

ROBERTSON, G.G., CARD, S.K., AND MACKINLAY, J.D. 1989. The Cognitive Coprocessor Architecture for Interactive User Interfaces. In Proc. User Interface Software and Technology (UIST '89), 10-18.

SARKAR, M. AND BROWN, M.H. 1992. Graphical fisheye views of graphs. In Proc. Human Factors in Computing Systems (CHI '92), 83-91.
SARKAR, M., SNIBBE, S.S., TVERSKY, O.J., AND REISS, S. P., 1993. Stretching the Rubber Sheet: A Metaphor for Viewing Large Layouts on Small Screens, In Proc. User Interface Software and Technology (UIST), 81-91.

SHNEIDERMAN, B. 1996. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In Proc. IEEE Symposium on Visual Languages.

WARE, C. 2000. Information Visualization: Perception for Design. Morgan Kaufman.

ZHANG, J. 1991. The Interaction of Internal and External Representations in a Problem Solving Task. In Proceedings of the Thirteenth Annual Conference of Cognitive Science Society, Lawrence Erlbaum Associates.