Effects of 2D Geometric Transformations on Visual Memory

Heidi Lam* Ronald A. Rensink* Tamara Munzner*
University of British Columbia

Abstract

We examined the effects of geometric transformations and their interactions with background grids on visual memory to provide interface design guidelines. We studied scaling, rotation, rectangular fisheye, and polar fisheye transformations. Based on response time and accuracy results, we defined a *no-cost zone* for each transformation type within which performance is unaffected. Results indicated that scaling had no effect down to at least 20% reduction. Rotation had a no-cost zone of up to 45 degrees, after which the response time increased to 5.4 s from the 3.4 s baseline without significant drop in accuracy. Interestingly, polar fisheye transformation had a lesser effect on accuracy than the rectangular counterpart. The presence of grids extended these zones and significantly improved accuracy in all but the polar fisheye transformation trials. Our results therefore provided guidance on the types and levels of nonlinear transformations that could be used without affecting performance, and provided insights into the roles of grids on visual memory and transformations.

CR Categories: H.1.2[Models and Principles]: User/Machine Systems—Human information processing H.5.2[Information Interfaces and Presentation]: User Interfaces—Evaluation/methodology

Keywords: Information visualization; visual memory; geometric transformations

1 Introduction

Geometric transformations are widely used in interface design, particularly in visualization systems where the amount of information to display exceeds available screen resolution, and in situations that require navigation through a two- or three-dimensional scene. For example, rotation has been used to navigate in network graphs displayed using a radial layout [Yee et al. 2001]. Scaling is extremely popular in interface design; for example, thumbnails are widely used as icons. Unfortunately, scaling only works to a certain extent. When the size of an image is reduced too far, its details become indiscernible. One possible remedy is to selectively scale the image such that readability can be preserved for the part of the image that is relevant to the user, while the rest remains available in a reduced form to serve as context for orientation. The class of Focus+Context techniques does so by providing both an unscaled focus and a scaled-down context in one integrated image [Leung and Apperley 1994, Skopik and Gutwin 2005]. Focus+Context can be realized using a nonlinear transformation called *fisheye* transformation, which has two main variants: rectangular and polar [Leung and Apperley 1994, Sarkar and Brown 1992]. There exists a large body of work using the fisheye transformation, including the Fisheye menu for text lists [Bederson 2000], calendar on small screens [Bederson et al. 2004] and 2D graph display for large information spaces [Bartram et al. 1995].

While scaling, rotation and fisheye transformations can provide benefits in visualizing information, there is a danger that the transformed image may be too distorted to remain recognizable. This issue is a serious usability concern, since users need to be able to retain, or at least compensate for, their orientation in the visualization after the transformation. They also need to be able to associate the components in the display before and after the transformation to equate the two views as the same, or at least holding the same information. Unfortunately, the effects of these transformations on human perception are largely unknown. Existing rule-of-thumb guidelines suggest ways to transform images with minimal disruption. These guidelines include:

- Maintain orthogonal ordering (left-right, up-down ordering), proximity (distance relationships between objects) and topology (inside-outside relationships) of the original image [Misue et al. 1995];
- Use visual cues to support the user's comprehension of geometric distortion [Carpendale et al. 1997]. Background grids have been suggested as the most effective of these [Zanella et al. 2002], as used in EPT [Carpendale et al. 1997].
- Use animation to retain the relationships among components displayed during transformation, and to avoid reassimilating the new display [Robertson et al. 1989]. Many visualizations involving geometric transformation follow this principle, with earlier adopters being Pad++ [Bederson and Hollan 1994] and Table Lens [Rao and Card 1994].

While these guidelines may provide designers with some hints for handling geometric transformations in an interface, they are based mostly on casual experience, and are not detailed or quantitative enough for actual implementation. Clearly, different types of geometric transformations and different degrees of transformation incur different amounts of perceptual costs. Knowing these costs would help designers gauge cost-benefit tradeoffs in their applications. Quantifying the effectiveness of various techniques suggested by these guidelines to mitigate transformation costs would be also helpful. For example, since smooth animation may impose a heavy computational load, it would be useful to quantify the largest degree of transformation "jump" we can perceptually tolerate. Also, the presence of grids may create visual noise instead of being beneficial.

Extending our earlier study on geometric transformations and visual search [Rensink 2004, Lau et al. 2004], the goal of this work is to better understand and quantify the effects of 2D geometric transformations on visual memory (VM) to guide interface and visualization design. In this paper, we present the first measurements of the effects of four types of geometric transformations on VM: scaling, rotation, rectangular fisheye, and polar fisheye transformations. These transformations were applied to automatically generated abstract images consisted of dots and connecting lines. We also investigated whether the common practice of using background grids helps performance [Zanella et al. 2002], or whether it impairs performance by introducing visual clutter.

---

*{hllam, rensink, tmm}@cs.ubc.ca

2 Related Work

Previous work has looked at the perceptual costs of geometric transformations in visual search tasks using abstract images. Rensink [2004] found no cost for translational shifts up to at least 2 degrees of visual angle, or 2 cm at a viewing distance of 55 cm. Performance was not significantly affected for rotations of up to 17 degrees, but degraded sharply beyond that. Scaling was found to be invariant at a reduction factor of 2, but created a measurable cost at 4. In another series of experiments involving visual search and nonlinear polar fisheye transformation, Lau et al. [2004] found that the transformation had a significant time cost, with performance slowed by a factor of almost 3 under large distortions. Interestingly, they did not find any benefits in adding grids to their images. In fact, grids caused performance to slow down, suggesting that they only added to the perceptual noise.

Several studies have examined the use of visual spatial memory in interface design. An example is Robertson et al.'s work on Data Mountain [1998]. However, relatively little appears to be known about on the effects of geometric transformations on visual memory. Skopik and Gutwin [2005] looked at the effects of rectangular fisheye transformation and found that distortions increased the time required to remember and find target nodes, but without affecting task accuracy.

3 Experiments

We conducted 10 original and 2 extended experiments to investigate the effects of geometric transformations on VM. All experiments used the within-subject design. In each experiment, we considered only a single factor, the transformation type, looking at five levels of transformation degree. Each transformation level consisted of two phases, each with 8 trials. In the learning phase, participants were presented with 8 stimuli in sequence. In the recognition phase, they were shown another set of 8 stimuli in sequence, 50% of which were shown in learning phase. For each stimulus, participants were asked to determine whether it had been shown in the learning phase. Baseline performance was measured in terms of response time and accuracy obtained using untransformed test stimuli. This baseline was then compared with results of the transformed trials.

3.1 Transformations

We applied four types of transformations to abstract images consisting of dots connected by lines: scaling, rotation, rectangular fisheye, and polar fisheye. We also examined the effects of grid presence and grid type. We initially carried out 10 experiments by applying the transformations to dot locations on the test stimuli:

- **Scaling** (1, 0.5, 0.33, 0.25, 0.2*x* reduction factor)
  Exp 1. no grid
  Exp 2. rectangular grid
- **Rotation** (0, 30, 45, 60, 90 degrees clockwise rotation)
  Exp 3. no grid
  Exp 4. rectangular grid
- **Rectangular fisheye** (0, 0.5, 1, 2, 3 transformation factor)
  Exp 5. no grid
  Exp 6. rectangular grid
  Exp 7. polar grid
- **Polar fisheye** (0, 0.5, 1, 2, 3 transformation factor)
  Exp 8. no grid
  Exp 9. rectangular grid
  Exp 10. polar grid

The transformation ranges were chosen based on two factors. For scaling, there was a physical limit to which we could reduce stimuli size without severely compromisingly perceivable details. Otherwise, we used pilot results to identify the start of performance degradation induced by the transformations.

Based on our results, we extended two of the experiments: (1) *experiment 4-ext*: rotation with a rectangular-grid to study a wider range of rotations: 0, 90, 120, 150, 180, and (2) *experiment 10-ext*: polar fisheye with a polar grid to study the effects of transforming the sizes of the dot, and drawing the connecting lines in various coordinate systems. We did not include the translation transformation as it had previously been found to be robust in visual search tasks to at least 2 degrees of visual angle [Rensink 2004].

3.2 Stimuli

All experimental stimuli were randomly generated abstract images consisting of dots connected by lines. We chose to use abstract rather than photorealistic images in part to avoid non-visual effects in the recognition, such as the verbal effect found by Goldstein and Chance [1971], where recognition accuracy was considerably lower for objects difficult to name. They found a 71% recognition rate for faces, almost double that of magnified snowflakes (33%), and inkblots (46%). Moreover, in the domain of information visualization, data is typically represented in abstract form. Our stimuli were similar to 2D network graphs, but we believe these results generalize to many abstract visual encodings of information.

All original stimuli had a resolution of 300x300 pixels to ensure all levels of transformations would fit onto the display screen. In all the grid experiments, we filled the entire screen with the corresponding grid. We used a different set of images for each experiment. All images were automatically generated in the same manner for consistency. Each image consisted of 15 dots connected by lines. The number of dots was determined in pilot studies to optimize image memorability. We found that when there were too many dots on the image, the image tended to be non-salient and therefore difficult to memorize, as the dots were typically squeezed into the allotted space as a tight matrix-like structure. However, an image with too few dots was so sparse that transformation effects could not be recognized. The locations of the dots were randomly generated. The algorithm only guaranteed non-collision but not constant density of the dots.

Pilot studies showed that the task was too difficult if we only provided the dots. Lines were therefore added to link the dots to enhance stimuli memorability, similar to lines drawn between stars in astronomical constellations. The algorithm that added the lines between the dots did not guarantee that all the dots were joined as a single unit, but it did ensure all of the dots were connected to at least one other dot, namely, its nearest neighbour. The algorithm minimized line crossing, but did not control the total number of topological features, for example loops.

When grids were added to the images, the thickness of the connecting lines was increased to 2 pixels to better distinguish the dot-line foreground from the grid background.

For the fisheye transformation experiments, we used a transformation function, taken from Leung & Apperley [1994]:

$$T(x) = \frac{(d+1)x}{(dx+1)} \tag{1}$$

where $T(x)$ is the transformed value given input x, and d is the transformation factor. A larger d value leads to a higher degree of distortion.

Figure 1 shows a series of stimuli showing all the transformation types and levels.

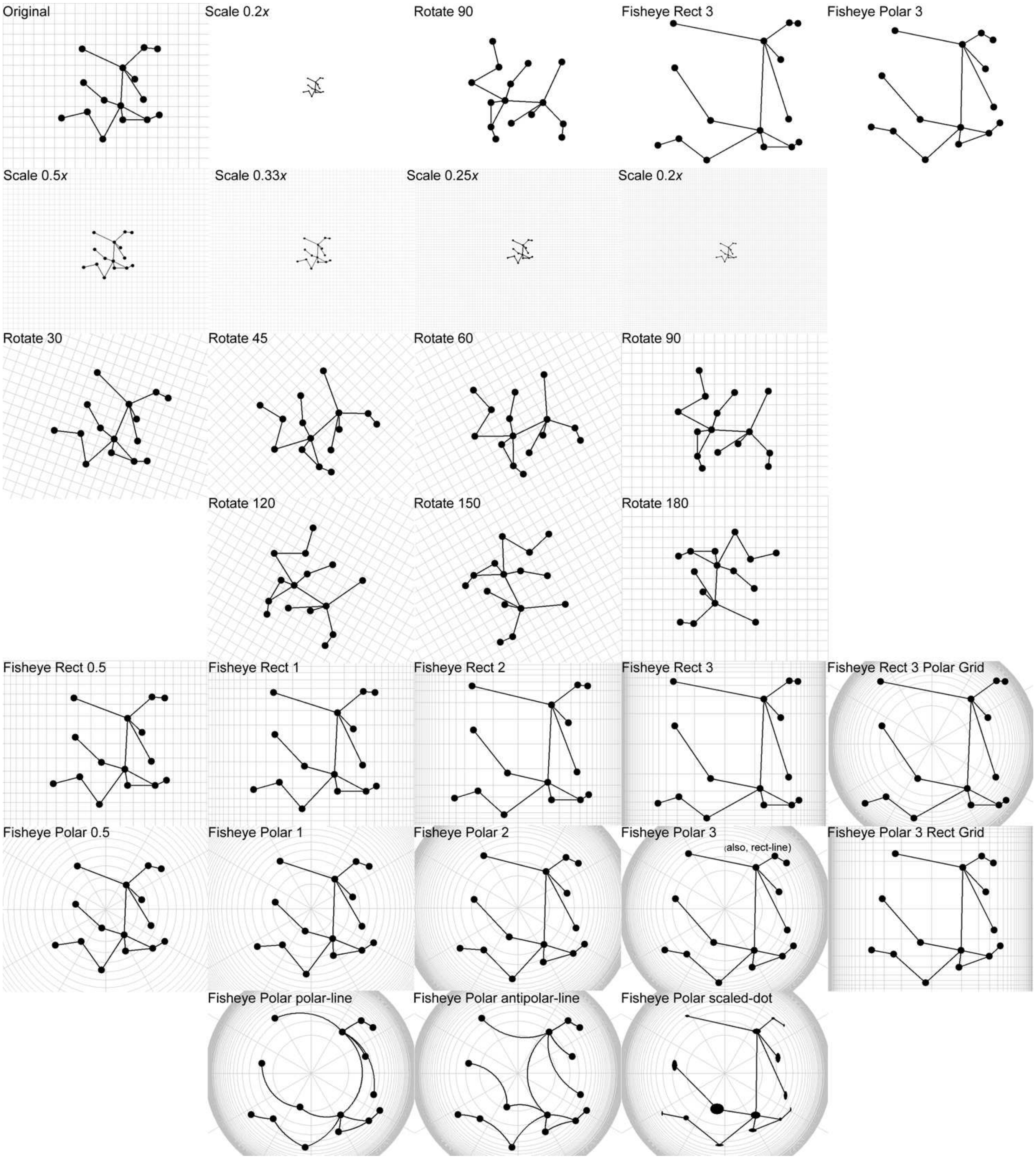

Figure 1: Sample stimuli showing transformation types and levels used in the 10 original and 2 extended experiments. The first row shows the maximally distorted stimuli used in the no-grid experiments (expts. 1, 3, 5, 8). The next 3 rows show all the distortion levels of the rectangular-grid experiments for scaling (expt. 2) and rotation (expt. 4; 4-ext). The fifth row shows all the distortion levels for rectangular-grid fisheye rectangular experiment (expt 6), along with the maximally distorted image for the polar-grid variety (expt 7). The last two rows contain the polar fisheye stimuli (polar-grid: expt 10 + rectangular-grid: expt 9; 10-ext).

3.3 Protocol

For each of the 12 experiments, all 20 participants completed trials on all five levels of the test transformation, and the order of appearance of the levels was counterbalanced between the participants. The stimulus was randomly selected from a pool of 50 and each only appeared once in the entire experiment to avoid learning effects. Prior to the actual experiment, participants were shown samples of original and transformed images to help them understand the transformation.

Each transformation-level session consisted of two phases: learning and recognition. In the learning phase, participants were asked to study 8 images one after another. Each was displayed for 12 seconds and followed by a 2.5-second blank screen before the next image appeared. Only untransformed images were used in the learning phase. Participants were told that they would need to recognize those images later on in the experiment, and that some of them might be transformed in a manner similar to the images shown prior to the experiment. In the recognition phase, 8 transformed images were shown to the participants in sequence. Half of these had been shown in the learning phase in their original form. The participants' task was therefore to indicate whether they had seen the images in the learning phase.

Prior to the actual experiment, participants were trained on the task using untransformed images in both the learning and the recognition phase. They were required to obtain at least 80% accuracy before starting the actual study.

The entire experiment typically took 30 minutes. Participants were compensated for their time with five dollars. In order to do well on the tasks, participants needed to pay close attention to the test images during the learning phase. As an added incentive, we informed the participants that high-accuracy scores would result in additional five-dollar bonuses.

3.4 Participants

A different group of 20 participants was tested in each of the 12 experiments. All were university students with normal or corrected-to-normal vision. Their ages ranged from 18 to 34 years.

3.5 Results and Data Analysis

We recorded two performance measurements: response time and accuracy. Response time was defined as the period from which the image was shown to the participant at the recognition phase, to the time when he responded with an answer. Accuracy was a percentage of answers that correctly identified whether the images had been shown in the learning phase. The accuracy baseline was 50%, as blind guessing would lead to chance, since half of the images shown in the recognition phase were present in the learning phase.

For the time measurement, we performed repeated measure single-factor Analysis of Variance (ANOVA) with transformation type as the factor for each experiment. We used the Greenhouse-Geisser adjustment and marked the results as *adjusted* if the sphericity assumptions were violated. Post-hoc analyses were performed for statistically significant results with Bonferroni correction and marked as *corrected*. For the accuracy results, we used the Friedman test for the initial analyses, and the Mann-Whitney test for post-hoc analyses. Only significant results were reported for the post-hoc analyses.

For each experiment, we mapped out a *no-cost zone* beyond which the performance began to degrade, as indicated by significantly higher response times and lower accuracy rates when compared to baselines established by trials using untransformed images.

Due to the large number of experiments, we first present a summary our results as a list of identified no-cost zones in Table 1. For cases where the boundaries were not indicated by statistical analyses, we provided estimates based on result trends, and marked them by a '?'. Table 2 lists the response time and accuracy results from trial at transformation levels immediately outside of the identified no-cost zones. Corresponding baseline values were provided in parentheses for comparison.

In summary, we found that VM was robust against many forms of transformations to a large extent. Scaling did not impact performance in terms of response time or accuracy down to a reduction factor of at least 0.2*x*. Rotation did not seem to affect performance up to 45 degrees, and both fisheye transformations had little effect on time or accuracy up to $d = 1$. We also found that the presence of grids extended these boundaries.

Table 1. Summary of experimental results: no-cost zones. No-cost zone indicates the largest transformations that can be compensated by our visual system without incurring a cost in performance. The combined result is the minimum of the time and accuracy results.

		No-cost zone		
	Experiment	**Time**	**Accuracy**	**Combined**
1.	Scaling: no-grid	*≥0.2x*	*≥0.2x*	*≥0.2x*
2.	Scaling: rect-grid	*≥0.2x*	*≥0.2x*	*≥0.2x*
3.	Rotation: no-grid	45°	45°?	45°
4-ext.	Rotation: rect-grid	60°	60°	60°
5.	Rect Fisheye: no-grid	$d = 1$	$d = 1$	$d = 1$
6.	Rect Fisheye: rect-grid	$d = 2$	$d = 2$	$d = 2$
7.	Rect Fisheye: polar-grid	$d = 2$?	$d = 2$	$d = 2$
8.	Polar Fisheye: no-grid	$d = 1$?	$d = 1$	$d = 1$
9.	Polar Fisheye: rect-grid	$d = 2$	$d = 2$	$d = 2$
10.	Polar Fisheye: polar-grid	$d = 2$	$d = 2$?	$d = 2$

Table 2. Summary of experiment results: performance cost as time and accuracy results at the next transformation levels just outside the no-cost zones, as shown in the Tx Level column. Baseline values were provided in parentheses for comparison. Italicized results were cases where the boundaries were estimated based on observed trends instead of statistical analyses.

			Performance Cost	
	Experiment	**Tx Level**	**Time (s)**	**Accuracy (%)**
1.	Scaling: no-grid	none	none	none
2.	Scaling: rect-grid	none	none	none
3.	Rotation: no-grid	60°	5.4 (3.4)	*69 (88)*
4-ext.	Rotation: rect-grid	90°	5.9 (4.1)	75 (88)
5.	Rect Fisheye: no-grid	$d = 2$	5.2 (4.6)	50 (88)
6.	Rect Fisheye: rect-grid	$d = 3$	3.9 (2.8)	75 (88)
7.	Rect Fisheye: polar-grid	$d = 3$	*5.5 (3.5)*	75 (94)
8.	Polar Fisheye: no-grid	$d = 2$	*4.7 (3.7)*	75 (94)
9.	Polar Fisheye: rect-grid	$d = 3$	5.6 (3.5)	75 (88)
10.	Polar Fisheye: polar-grid	$d = 3$	5.6 (3.8)	*75 (88)*

Detailed Results and Statistics

We now provide the detailed experimental results and data analyses for each of the four transformation types.

3.5.1 Transformation Type 1: Scaling

The scaling time and accuracy results showed no significant differences between the five levels, with or without adding grids to the images: time/no-grid: F(2.3, 43.2) = 0.67, *p* = .54, *adjusted*; score/no-grid: $\chi^2(4, N=20) = 2.01$; time/rect-grid: F(4, 76) = .60, *p*

= .67; score/rect-grid: $\chi^2(4, N=20) = 3.15$, $p = .53$. This result indicated that the range of scaling we studied did not impact performance.

3.5.2 *Transformation Type 2: Rotation*

Figure 2 shows the results. For the no-grid experiment, we found a marginal main effect in time (F(1.9, 35.8) = 2.92, $p = .070$). Post-hoc analysis indicated that time performance began to degrade at 60 degree, at which participants were slower at 5.4 s compared to the 3.4 s baseline. We also found a marginal main effect in accuracy ($\chi^2(4, N=20) = 8.75$, $p = .070$) but could not identify a clear no-cost accuracy boundary.

For the rectangular-grid experiment, we failed to find a main effect in both time (F(2.6, 49.7) = 1.33; $p = .27$, *adjusted*) and accuracy ($\chi^2(4, N=20) = 7.16$, $p = .13$), thus we were unable to locate no-cost zone boundaries based on these results.

Since we found relatively little performance degradations in the rectangular-grid results, we extended the range of rotation studied in another experiment to cover 0, 90, 120, 150 and 180 degrees (experiment 4-ext). The results are shown in Figure 4 as "Rectangular Grid Ext".

We found a main effect in time (F(4, 76) = 5.05, $p = .001$). Post-hoc analysis indicated both the 90-degree and the 180-degree rotation trials were significantly slower at 5.9 s compared to the 4.1 s baseline. We also found a main effect in accuracy ($\chi^2(4, N=20) = 14.95$, $p = .005$). Post-hoc analysis indicated the transformed trials were 14% less accurate than baseline. These results therefore suggested a no-cost boundary of 60 degrees.

In addition to no-cost zone boundaries, we were also interested in the performance improvement provided by the rectangular grid. We thus compared the accuracy between the non-grid and grid trials from 30 to 90 degree rotations, and found that the grid results were higher than their non-grid counterpart by 10% (two-tailed Mann Whitney test, $p = .03$). This increase in accuracy was not accompanied by an increase in time, thus ruling out any time-accuracy tradeoff.

3.5.3 *Transformation Type 3: Rectangular Fisheye*

Figure 3 shows the results. For the no-grid experiment, we found a marginal main effect in time (F(1.9, 36.2) = 2.83, $p = .074$, *adjusted*). It took 0.6 s longer for $d = 2$ and $d = 3$ trials than the 4.6 s baseline. We also found a main effect in accuracy ($\chi^2(4, N=20) = 43.80$, $p < .001$) and the $d = 2$ and $d = 3$ trials were 33% less accurate than the rest of the trials. Using the one-sample z-test, we found that the accuracy for $d = 2$ and $d = 3$ trials were at chance (Z(N=40) = 1.44; $p = .15$). These results indicated a clear no-cost zone boundary at $d = 1$.

For the rectangular-grid experiment, we found a marginal main effect in time (F(2.78, 52.9) = 2.63; $p = .063$, *adjusted*). Post-hoc analysis indicated that $d = 3$ trials were slower at 3.9 s when compared to the 2.8 s baseline, indicating a no-cost time boundary at $d = 2$. There was a strong effect in accuracy ($\chi^2(4, N=20) = 18.34$, $p = .001$), with baseline and $d = 1$ trials being 15% more accurate from those of $d = 3$, indicating a no-cost accuracy boundary at $d = 2$.

For the polar-grid experiment, the main effect in time was also marginal (F(4, 68) = 3..32; $p = .051$, *adjusted*), with a marginal time degradation at $d = 3$ ($p = .077$, *corrected*). While the task accuracy main effect remained, it was much smaller ($\chi^2(4, N=19) = 10.4$, $p = .034$), with a no-cost accuracy boundary at $d = 2$.

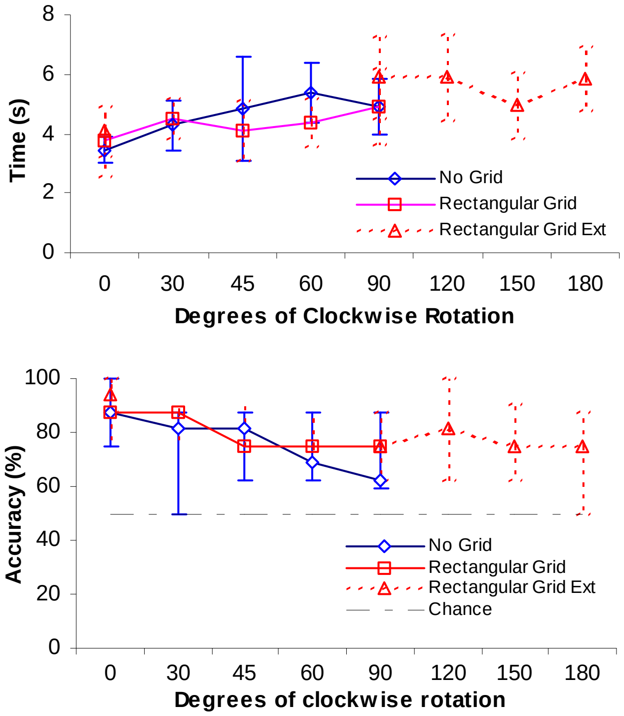

Figure 2. Results for the rotation experiments. Time data points are averages with 95% confidence interval bars. Accuracy results are medians with quartiles.

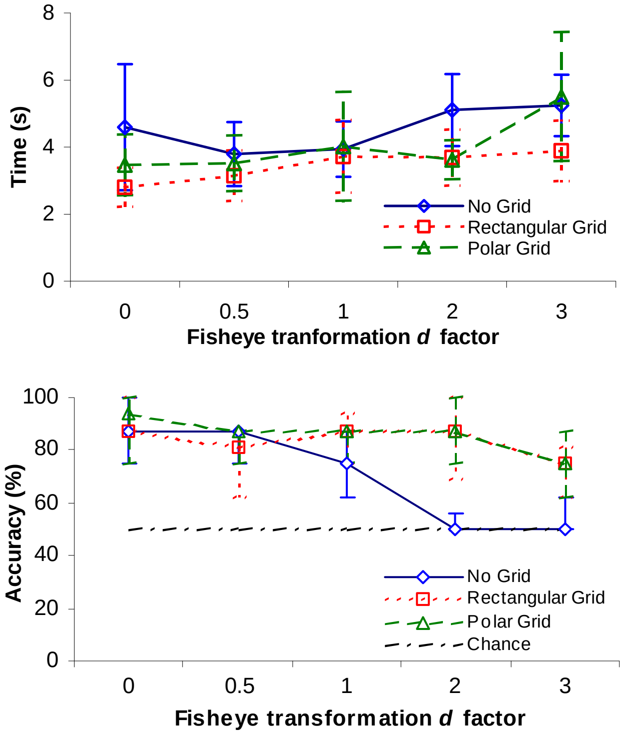

Figure 3. Results for the rectangular fisheye experiments. Time data points are averages with 95% confidence interval bars. Accuracy results are medians with quartiles.

3.5.4 Transformation Type 4: Polar Fisheye

Figure 4 shows the results. We failed to find a main effect in time for the no-grid experiment (F(1.82, 34.5)=2.3; p = .12, *adjusted*). There was, however, a main effect in accuracy (χ^2(4, N=20) = 17.16, p = .002), with d = 2 and d = 3 trials being 20% less accurate than baseline, thus indicating a no-cost accuracy boundary at d = 1. One-sample z-test indicated that performance at d = 2 and d = 3 had not degraded to chance (Z(N=40) = 8.23; p < .001).

For the polar-grid experiment, we found a main effect in time (F(4, 76) = 6.08, p = <.001). Post-hoc analysis indicated d = 3 trials were 1.7 s slower than baseline and d = 1 trials, which took 4 s on average. This indicated a time no-cost zone boundary at d = 2. We failed to find a main effect in accuracy (χ^2(4, N=20) = 6.92, p = .14).

For the rectangular-grid experiment, we found a main effect in time (F(4, 76) = 4.32, p = .003). Post-hoc analysis indicated d = 3 trials were slower by 1.8 s than the 3.8 s baseline and d = 1 trials, thus indicating a no-cost time boundary at d = 2. We also found an accuracy main effect (χ^2(4, N=20) = 11.27, p = .024). Post-hoc analysis indicated d = 3 trials were 12% less accurate than baseline, thus indicating a no-cost accuracy boundary at d = 2.

Despite extending the no-cost boundaries from d = 1 to 2, the presence of either polar or rectangular grids on polar fisheye transformed images did not substantially improve accuracy. This result was in stark contrast to that found in the rectangular fisheye experiments. Given the relatively high accuracy in the no-grid polar trials outside of the no-cost zone, we wondered if we have reached the accuracy ceiling, and grids could not further enhance VM. In our previous experiments, we drew straight lines between the dots. This may not be as natural in the polar transformed images as in their rectangular counterparts. We wondered if this unnaturalness may contribute to our observed lack of benefit of grids in the polar trials, or we had truly reached the limit of VM compensation.

To test our hypothesis, we extended the polar fisheye experiment to look at line shape (experiment 10-ext), where the straight lines in the original images were drawn based on the polar coordinate system (*polar-line*), the rectangular coordinate system (*rect-line*), or a mirror image of the ones drawn in the polar coordinate system (*antipolar-line*). The last case was included to tease out any potential adverse effects induced by an extremely unnatural transformation on the lines. We also included a case where we transformed the size of the dots and keeping the lines in the rectangular coordinate system (*scaled-dot*) to test if the dot sizes should themselves be transformed. Theoretically, transformation can be applied globally to the surrounding space, or locally to the objects in the space. Our earlier experiment assumed the model of transforming the space without affecting the sizes or shapes of the dots and the lines, as if they were merely pinned on the surface instead of completely adhered to the surface of transformation. The only exception was in scaling, where we had to transform the dot size to avoid collision.

We failed to find a main effect in time (F(2.4, 45.5) = 2.09, p = .13), but we did find a main effect in accuracy (χ^2(4, N=20) = 15.7, p = .003). Post-hoc analysis indicated that our participants made significantly more errors in the *polar-line* trials than baseline, and the accuracy was at chance (Z(N=20) = 1.45; p = .15).

Examples of these transformations are shown in the last row of Figure 1 and Figure 5 shows the results.

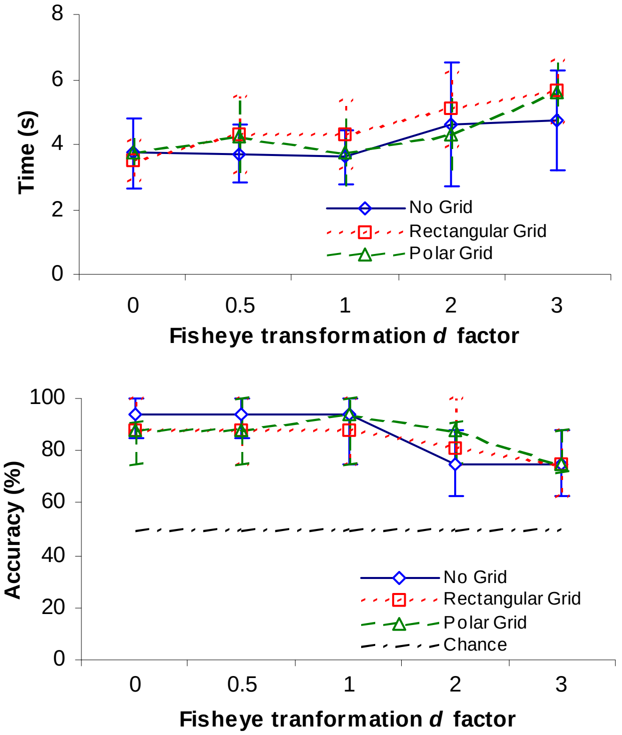

Figure 4. Results for the polar fisheye experiments (d effect). Time data points are averages with 95% confidence interval bars. Accuracy results are medians with quartiles.

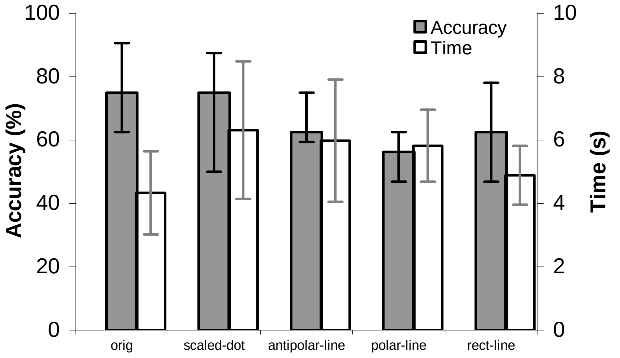

Figure 5. Results for the extended polar fisheye experiment. Time data points are averages with 95% confidence interval bars. Accuracy results are medians with quartiles. orig = original image; scaled-dot = dot sizes transformed; antipolar-line = lines drawn as the mirror image in the polar coordinate system; polar-line = lines drawn in the polar coordinate system; rect-line = lines drawn in the rectangular coordinate system.

4 DISCUSSION

Our experimental results mapped no-cost zones in all the transformation types where VM was not significantly disrupted. We also found significant benefits of grids in all transformation types studied except in our polar fisheye trials.

4.1 Effects of Image Transformations

This work is part of a series of investigations to map out the extent to which users can cope with geometric transformations in visual interfaces. Our current results and those reported previously suggested that invariance was possible for all geometric transformations for up to a point. Interestingly, this invariance appeared to be more extensive in recognition than search tasks. For example, search task performance degraded after a 50% reduction, while memory task performance remained unaffected even at 20% of the original size. Similarly, our participants could tolerate a larger distortion in rotation (memory: 45°; search: 17°), and polar fisheye transformation (memory: d = 1; search: d = 0.5)[+].

While we applied the transformations to dot locations in most of our experiments, we found interesting results when we applied the polar fisheye transformation to dot sizes, and drew the connecting lines based on different coordinate systems. Contrary to our intuition, trials using images with lines drawn based on the polar coordinate system were least accurate and at chance, while corresponding trials with supposedly unnatural mirror images of these lines exhibit better performance. These results seemed to suggest that image distinctiveness, rather than component consistency, was a more important factor in memorability. At large distortions, the lines in the *polar-line* images formed similarly rounded shapes, while the corresponding *antipolar-line* images produced figures with enough acute angles to remain distinguishable, despite their blatant incongruity with the underlying transformation and with the coordinate system.

4.2 Effects of Grids

Adding grids to the images appeared to help in two ways:

1. *No-cost zone extension*: The presence of either rectangular or polar grid pushed the no-cost zone boundaries to higher levels. For example, the combined no-cost zone boundary for the fisheye transformations were increased from d = 1 to d = 2, and the rotation boundary was pushed from 45 to 60 degrees.
2. *Accuracy improvement*: Grids were found to improve accuracy. For rotation, participants were 10% more accurate in grid trials without spending extra time in the task, thus ruling out potential time-accuracy tradeoffs. In the case of rectangular fisheye transformation, we found that participants' accuracy improved from chance to baseline at d = 2, and to 75% at d = 3, again without time compensation. Interestingly, we failed to observe substantial improvement by adding grids to polar fisheye transformed images.

To understand the apparent lack of performance improvement in polar trials, and to obtain further insights to the different transformation types and their interactions with grids, we revisited the design guidelines described in the *Introduction* section.

4.3 Revisiting design guidelines

Misue et al. [1995] suggested that horizontal/vertical ordering, proximity and topology should be maintained to minimize disruptions incurred by image transformations. Indeed, the scaling transformation preserves all three, and the limit of transformation seems to be how far can one reduce the image physically before the details can no longer be perceived. This finding is consistent with the common interface design practice of using scaled-down versions of images to represent full-resolution file contents, especially when the file content is visually salient, as in the cases of most image files and graphically intense web pages. Indeed, various forms of thumbnails have been suggested for small-screen devices to avoid the laborious reauthoring of desktop-sized web pages for small screens, such as textually-enhanced thumbnails [Woodruff et al 2001] and WebThumb [Wobbrock et al 2002].

The rotation transformation violates horizontal/vertical ordering but still maintains proximity and topology. Interestingly, the rectangular grid fails to benefit performance starting at 90-degree rotation. Since our images did not have a clear up-down axis, this limit may be due to our inability to recognize the main vertical axis and the up direction in the image. Having a rectangular grid may help re-orientation, but only to a certain extent as the information provided by the grid became ambiguous starting from a rotation angle of 90 degrees. For example, the grid looked the same for 0, 90 or 180-degree rotations, and similarly for 30 or 120, and 60 or 150-degree rotations. Taken together, our results suggested a refinement to Misue et al.'s guideline on maintaining orthogonal ordering: transformation should preserve an orthogonal relationship between principal axes with a clear up-down axis.

For both fisheye transformations, the proximity relationship is violated while preserving horizontal/vertical ordering and topology. In that case, the perceptual challenge is to discern the relative distance between objects in the image. The polar fisheye transformation seemed to be much better tolerated than its rectangular counterpart, as accuracy was maintained at 75% even outside the no-cost zone in the polar case while corresponding rectangular trials showed chance performance. This result was not expected, as the polar transformation's rounded appearance does not look natural on a rectangular screen [Leung and Apperley 1994]. In addition, it bends horizontal and vertical lines. Nonetheless, polar fisheye transformation is generally preferred over its rectangular counterpart in map applications, since the distortion may be perceived as consistent to the effect of distorting the map onto a hemisphere, and the transformation preserves the angle of the original image [Sarkar and Brown 1992, Churcher et al. 1997]. The polar fisheye transformation may also be more familiar than rectangular, as the effect resembles that produced by the ultra-wide angle fisheye lens used in photography.

The number of transformation parameters and their degree of integration in the transformation may further explain the lesser performance degradation observed in our polar fisheye trials. In the rectangular case, the width and the height are transformed separately. Rectangles that are of the same distance from the focus point may not have the same size and shape. As a result, objects may be distorted with different aspect ratios based on their horizontal and vertical distances, which may impose a higher mental load [Bartram et al. 1995]. In contrast, the polar fisheye transformation only distorts radial distances, and may not carry the same problem as in the rectangular case.

This issue may also explain the different effect that we observed in our fisheye transformation trials. In the rectangular fisheye trials, addition of either a polar or rectangular grid improved accuracy from chance to 75% without time compensation. In contrast, neither a rectangular nor a polar grid improved performance in the corresponding polar fisheye trials. One possible explanation is that the grid, may it be polar or rectangular, provided a powerful visual cue encoding standard distances in transformed images that may help to offset the difficulty in distance estimation when the image is distorted, as in the rectangular fisheye case. Since distance transformation is integrated in polar fisheye transformations, distance estimation may not be as difficult as in the rectangular case, thus nullifying potential benefits brought about by adding a grid.

[+] The [Lau et al. 2004] experiments used a different fisheye polar transformation function with a transformation factor *c*. A *c* value of 1.2 can be roughly translated to our d = 0.5. The fisheye rectangular transformation was not tested in their experiments.

Smooth animation is another technique believed to alleviate the disruptive effects of image transformations [Robertson et al. 1989, Bederson and Boltman 1999]. Similar to our earlier work on visual search [Rensink 2004, Lau et al., 2004], our current results suggested the visual system could compensate for relatively large ranges of transformation distortions. Both visual search and visual memory have thus been ruled out as valid reasons for a smooth animation requirement. Nevertheless, the need for smooth animation may arise from some other perceptual mechanisms, so we would need further investigations before advocating relaxing that design guideline.

5 Conclusions

We examined the effects of four different types of transformations on visual spatial memory: scaling, rotation, rectangular fisheye, and polar fisheye transformations. We found no-cost zones in all of the transformation types that far exceed those found in our earlier work on visual search. We also found substantial benefits in applying grids to images for all of our transformation types except for the polar fisheye case. Our work therefore quantified the limits of our visual spatial memory in coping with geometric transformations, and validated the use of grids as a visual cue to aid recognition of images.

We adopted the view that geometric transformation was considered to simply affect the location of objects within a space in this work. An equally valid view is to consider the transformation on the space itself, including the objects embedded within it. Translating that view onto our current experimental set up, we would need to transform dot sizes and line shapes in addition to dot locations. This is a subtle but important distinction, since in the latter case, visual cues providing more information about how the space has transformed could improve performance. We started to study space transformation in our extended study on the polar transformation, where we looked at the effects of transforming dot sizes, and drawing the connecting lines based on either the polar or the rectangular coordinate system. Our results indicated that memorability is more contingent upon image distinctiveness rather than consistency with the underlying transformation and coordinate system. Further investigations are warranted to better quantify the differences in the two views.

Our experiments looked at how a single and uniform transformation affects visual memory. In real-life situations, images may transform by parts and independently. It would be interesting to compare our results with those obtained using multiple transformations on a single image. We suspect the perceptual limits for multiple transformations will be much smaller than those established in our current set of experiments. It is also interesting to look at single or multiple transformation types with single or multiple degree of transformation.

The question of scalability is also of interest. We decided on a small number of dots in the stimuli to create an acceptable level of task difficulty. It would be interesting to see if the total number of dots in the stimuli would impact visual spatial memory in similar ways if the stimuli contain components that are individually salient and memorable.

6 ACKNOWLEDGMENTS

We would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding, Minjung Kim and Raman Singh for their assistance in running the experiments, and Melanie Tory, Daniel Archambault, Peter McLachlan and James Slack for providing feedback on the paper draft.

7 REFERENCES

BARTRAM, L., HO, A., DILL, J., AND HENIGMAN, F. 1995. The Continuous Zoom: A Constrained Fisheye Technique for Viewing and Navigating Large Information Space. In *Proc ACM UIST*, 207-215.

BEDERSON, B.B., HOLLAN, J.D. 1994. Pad++: A Zooming Graphical Interface for Exploring Alternate Interface Physics. In *Proc ACM UIST,* 17-26.

BEDERSON, B., AND BOLTMAN, A. 1999. Does animation help users build mental maps of spatial information. In *Proc IEEE InfoVis*, 28-35.

BEDERSON, B. 2000. Fisheye Menus. In *Proc ACM UIST*, 217-225.

BEDERSON, B., CLAMAGE, A. CZERWINSKI, M., ROBERTSON, G.G. 2004. DateLens: A fisheye calendar interface for PDAs. *ACM ToCHI,* 11(1), 90-119.

CARPENDALE, M.S.T., COWPERTHWAITE, D.J. AND FRACCHIA, F.D. 1997. Making distortions comprehensible. In *Proc IEEE Symposium on Visual Languages*, 36-45.

CHURCHER N., PRACHUABMOH, P. AND CHURCHER, C. 1997. Visualization techniques for collaborative GIS browsers. *International Conference on GeoComputation.*

GOLDSTEIN, A.G. AND CHANCE, J.E. 1971. Visual Recognition Memory for Complex Configurations. *Perception and Psychophysics* 9(2), 237-241.

LAU, K., RENSINK, R.A., AND MUNZNER, T. 2004. Perceptual Invariance of Nonlinear Focus+Context Transformations. In *Proc ACM APGV*, 65-72.

LEUNG, Y.K. AND APPERLEY, M.D. 1994. A Review and Taxonomy of Distortion-Oriented Presentation Techniques. *ACM ToCHI*, 1(2), 126-160.

MISUE, K., & EADES, P., & LAI, W. & SUGIYAMA, K. 1995. Layout adjustment and the mental map. *Journal of Visual Languages and Computing*, 6(2) 183-210.

RAO, R. AND CARD, S. 1994. Table lens: Merging graphical and symbolic representations in an interactive focus plus context visualization for tabular information. In *Proc ACM CHI,* 318-322.

RENSINK, R.A. 2004. The Invariance of Visual Search to Geometric Transformation. *Journal of Vision 4*.

ROBERTSON, G.G., CARD, S.K., AND MACKINLAY, J.D. 1989. The cognitive coprocessor architecture for interactive user interfaces. In *Proc. ACM UIST*, 10-18.

ROBERTSON, G., CZERWINSKI, M., LARSON, K., ROBBINS, D., THIEL, D., AND VAN DANTZICH, M. 1998. Data Mountain: using spatial memory for document management. In *Proc ACM UIST*, 153-162.

SARKAR, M. AND BROWN, M.H. 1992. Graphical fisheye views of graphs. In *Proc ACM CHI,* 83-91.

SKOPIK, A., AND GUTWIN, C. 2005. Improving revisitation in fisheye views with visit wear. In *Proc ACM CHI*, 771-780.

WOODRUFF, A., FAULRING, A., ROSENHOLTZ, R., MORRISON, J., AND PIROLLI, P. 2001. Using Thumbnails to Search the Web. In *Proc ACM CHI*, 198-205.

WOBBROCK, J., FORLIZZI, J., HUDSON, S., MYERS, B. WebThumb: Interaction Techniques for Small-Screen Browsers. 2002. In *Proc ACM UIST,* 205-208.

YEE, K.P., FISHER, D., DHAMIJA, R., HEARST, M. 2001. Animated Exploration of Graphs with Radial Layout. In *Proc IEEE InfoVis*, 45-50.

ZANELLA, A., CARPENDALE, M.S.T., ROUNDING, M. 2002. On the effects of viewing cues in comprehending distortions. In *Proc ACM CHI,* 119-128.