Applying guidelines for the design of distortions on focus+context interfaces

Simon Butscher
Human-Computer Interaction Group,
University of Konstanz
Simon.Butscher@uni-konstanz.de

Harald Reiterer
Human-Computer Interaction Group,
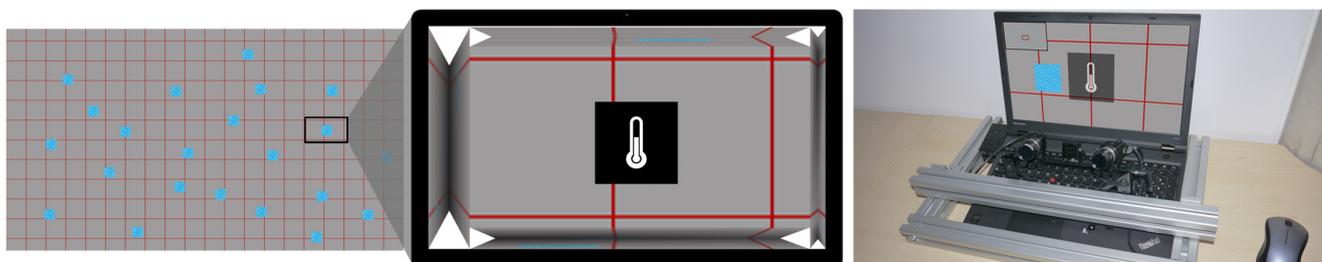University of Konstanz
Harald.Reiterer@uni-konstanz.de

Figure 1. (left) Mutliscale space containing several veiled objects; (middle) FrameFold: Multiscale space is folded on the edges of the viewport; (right) Evaluation setting showing the overview condition.

ABSTRACT
Distortion-based visualization techniques allow users to examine focused regions of a multiscale space at high scales but preserve their contextual information. However, the distortion can come at the coast of confusion, disorientation and impairment of the users' spatial memory. Yet, how distortions influence users' ability to build up spatial memory, while taking into account human skills of perception, interpretation and comprehension, remains underexplored. This note reports findings of an experimental comparison between a distortion-based focus+context interface and an undistorted overview+detail interface. The focus+context technique follows guidelines for the design of comprehensible distortions: make use of real-world metaphors, visual clues like shading, smooth transitions and scaled-only focus regions. The results show that the focus+context technique designed following these guidelines help to keep track of the position within the multiscale space and does not impair users' spatial memory.

CCS Concepts
• **Human-centered computing~Visualization techniques**
• **Human-centered computing~Graphical user interfaces**

Keywords
Spatial memory; multiscale space; distortion; design guidelines; focus+context; overview+detail

1. INTRODUCTION
Multiscale information offers an overview of data at small scales and details at larger scales. Distortion-based visualization techniques, such as fisheye views [5] or folding techniques like

DOI: http://dx.doi.org/10.1145/2909132.2909284

mélange [7], allow users to examine focused regions at high scales but at the same time preserve their contextual information. However, the distortion can come at the cost of confusion, disorientation [4] and impairment of the users spatial location memory [15].

Although, spatial memory plays an important role in graphical user interfaces, research on the influence of focus+context interfaces on users' ability to memorize object locations is limited. In this context fisheye distortions are the only focus+context technique which has been investigated. Comparisons of fisheye visualizations with non-distorted interfaces showed a negative effect of distortions on the memorability of object locations and claims were made that distortions hinders the users from building up spatial memory. However, how distortions influence users' ability to build up spatial memory, while taking into account human skills of perception, interpretation and comprehension, remains underexplored.

This note reports findings of an experimental comparison between a distortion-based focus+context interface and an undistorted interface using an overview+detail design pattern. The distortion of the focus+context technique follows guidelines for the design of comprehensible distortions derived from findings of Carpendale et al. [4]: make use of real-world metaphors and visual clues like shading, smooth transitions and scaled-only focus regions. The distortion-based focus+context technique and the undistorted overview+detail interface are compared in terms of spatial memory and attention patterns.

2. SPATIAL MEMORY IN MULTISCALE SPACES
Knowledge about the interface layout and object locations can substantially reduce the cognitive and physical effort required for interaction [14]. Spatial consistency strongly influences the users' ability to build up spatial location memory. In this context consistency can be defined as the stable location of objects in relation to a frame of reference like the borders of a window or the display itself. In multiscale spaces this consistency is violated as the position of objects change based on the zoom level or the position of the viewport while panning.

The overview+detail design pattern tries to overcome this problem by providing an overview in a distinct area. This helps to anchor

the detailed view to the absolute spatial position in the multiscale space and enables the encoding of object locations in relation to the borders of the space as a frame of reference. Different studies for mobile devices showed that the presence of an overview increases spatial recall accuracy [2] and facilitates incidental learning of spatial locations [11]. In contrast to these results Hornbæk et al. [8] found no evidence for improved spatial memory with an overview compared to a multiscale space without an overview.

Instead of showing an overview, focus+context techniques integrate detailed information in their context. Distortion-based focus+context techniques allow for a magnification of focus regions while the border of the multiscale space as a frame of reference remains visible. Fisheye views are a well-studied example of such focus+context techniques [5]. However, a potential problem with distortion is the misinterpretation of the underlying data. Using distortion can come at the cost of inconsistencies of object locations and therefore can handicap the user in making spatial judgments which causes difficulties in remembering object locations [4]. Yet, fisheye views are the only focus+context approach which has been explored in terms of their effect on spatial memory [15]. Furthermore, previous research about the influence of focus+context distortions on users' spatial memory did not take users' abilities to interpret distortions into consideration.

3. DESIGN GUIDLINES FOR DISTORTION-BASED INTERFACES

Carpendale et al. [4] summarized findings about users' abilities to interpret distorted spaces. We extracted three guidelines which are applicable to focus+context interfaces:

(G1) *Scaled-only focus region*: Pure magnifications are easier to interpret than distortions. Therefore in focus regions only the scale factor should change and the regions should be smoothly integrated in the distorted context [13].

(G2) *Real-world metaphors*: 3-dimensional distortions (e.g. [7,9,10]) based on real-world metaphors provide rich visual information about how the space is distorted and therefore allow the user to quickly and accurately interpret the distortion. A combination of a perspective view and shading helps the user to identify the 3-dimensional shape the distortion is based on [12].

(G3) *Smooth transitions*: Many studies report on the need for continuous visual transitions (e.g. [3]). Seeing how the distortion is created or being able to reverse the transition allows the user to decode the distortion.

These guidelines can help design distortion-based interfaces which "leave the user confident the information they are exploring remains consistent" [4]. However, it is still underexplored how focus+context interfaces designed following these guidelines influence users' ability to build up spatial memory. Furthermore, these distortions imply a visual separation between the scaled-only focus regions and the distorted context, which is important to interpret the distortion. However, this visual separation can induce the problem of split attention, like it is already known for overview+detail interfaces. Our research questions therefore are:

(R1) Does a distortion-based focus+context interface designed following the guidelines for the creation of comprehensible distortions improve users' spatial memory?

(R2) Does the distorted context catch user's attention and if so, does a visual separation between a scaled-only focus region and its distorted context induce problems of split attention?

4. EVALUATION

To answer these research questions we conducted a comparative experiment between a distortion-based interface and an undistorted interface. We evaluated two visualization techniques for multiscale spaces, where both provide a frame of reference that can be used to anchor the viewport within the multiscale space and therefore help the user to build up spatial memory. The first condition uses an overview window to show a frame of reference when zoomed in. The second condition is a focus+context technique which folds the multiscale space at the edges of the display in order to provide clues about the distance to the border of the space.

Overview condition: In the *Overview* condition an overview+detail design pattern was used (see Figure 1, right). The overview contained veiled objects and a red rectangle that illustrates the currently visible viewport. The overview can help to encode object positions in terms of the distances to the border of the multiscale space as frame of reference. Whereas this condition provides a persistent and undistorted overview over the multiscale space it can lead to split attention in terms of switching between the detail and the overview area which requires mental effort and time [8].

FrameFold condition: Like other techniques [7,9] *FrameFold* is a focus+context interface which is based on distortions using the metaphor of a folded sheet of paper. In contrast to other techniques, it is designed as single focus technique which is optimized to fit the whole multiscale space within the visible viewport while panning and zooming. Instead of pushing parts of the multiscale space outside of the visible viewport *FrameFold* allows for scaling the focus region while folding its context **(G1)**. Folds are automatically created on the edges when the scaled space exceeds the size of the viewport (see Figure 1, middle). The width, the depth and the shading of a fold corresponds to the space that is aggregated within a fold **(G2)**. The folds give a good awareness of the space that is bridged [7] and therefore can help to encode the distance of an object to the border of the multiscale space as a frame of reference. The corners of the viewport can be intersected by both a horizontal and vertical fold. This leads to nonlinear distortions in this areas which are more difficult to interpret. In order to assist in the perception of the distance to the edge of the multiscale space, we avoided overlaps of the horizontal and vertical folds (see Figure 2, red circles). Furthermore the 3-dimensional geometry of the folds is corrected with respect to a virtual camera so that the valley of the fold is visible and not overlapped due to perspective distortion (see Figure 2). During panning and zooming the folds are instantly adapted to bridge the intervening space to the border of the multiscale space. This allows for smooth transition between a distorted view while zoomed in and an undistorted view while zoomed out **(G3)**.
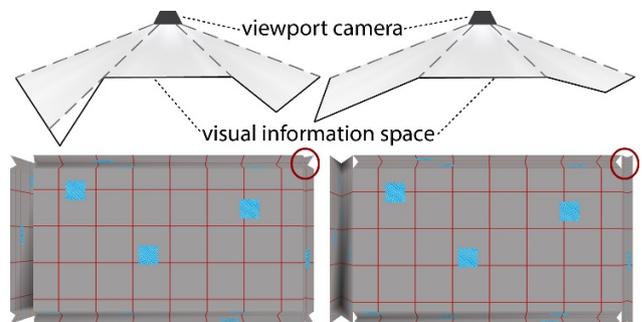

Figure 2. Perspective correction to make the valleys of the folds visible. (left) uncorrected view; (right) corrected view

4.1 Tasks

The participants had to perform two different tasks: a navigation task and reconstruction task based on it. Within the navigation task participants had to successively find eight objects placed on a multiscale space (1920*1080 pixels in size, see Figure 1, left). To force the users to navigate the multiscale space the objects were unveiled on a certain zoom level (1:8 scale). Each object to be found was visualized in the center of the screen (see Figure 1, right). The participants had to bring the corresponding object on the multiscale space to an overlap with the template in the middle of the screen. For the reconstruction task participants were instructed to recall the positions of the eight objects and place them on the empty multiscale space as accurately as possible. During the reconstruction task it was still possible to pan and zoom.

4.2 Design

The study was conducted as a within-subjects factorial design with two independent variables: interface (*FrameFold*, *Overview*) and object pool (Pool A, Pool B). In each interface condition participants used a separate object pool with different symbols and positions. An object pool consisted of 8 objects which had to be found and 14 objects which served as distractors to increase the difficulty of the task. As learning the location of objects is a natural side-effect of interacting with them [6] the navigation task had to be performed six times in succession. This results in 48 trials per participant (8 objects x 6 blocks). Each block started at a 1:1 scale, which allowed the entire space to fit within the visible viewport. The order of the objects within a block was altered to prevent the participants' remembering the navigation steps instead of learning the object locations. After the 6 blocks, participants performed the reconstruction task using the same interface as for the navigation task. Before each condition, a training task with a training object pool for the navigation as well as for the reconstruction was conducted. The study was systematically counterbalanced and each participant was randomly assigned to one of the resulting four groups (combination of technique order and pool assignment).

Dependent variable for the navigation task was the task completion time. Furthermore we used an eye tracker to investigate attention patterns. For the reconstruction task we calculated the recall error rate as Euclidian distance in pixels between the original location and the participants' recalled location. We did not record time for the reconstruction task as participants were instructed to take as much time as the need to place the objects.

The experiment was performed on a notebook sized display (15.4") with a resolution of 1920 x 1080 pixels (see Figure 1, right). Panning and zooming the multiscale space was done via mouse input. We recruited 12 participants (6 female) for the study. Participants were between 23 and 39 years old (M = 25.9, SD = 4.8). Only one participant had a background in computer science. In total we collected data from 1152 trials for the navigation task (12 participants × 2 interfaces × 8 objects × 6 blocks) and 192 trials for the reconstruction task (12 participants × 2 interfaces × 8 objects).

4.3 Results

The reporting of study results is structured into the navigation task and the reconstruction task. We used non-parametric tests to analyze the data, as the assumption of normality was violated.

4.3.1 Navigation Task

The mean time to find an object in the *Overview* condition (M = 19.6 sec, SD = 30.6) was slightly lower than in the *FrameFold* condition (M = 21.3 sec, SD = 24.2; see Figure 3, left). However the Wilcoxon signed rank test revealed that this difference was statistically not significant (p > .05).

The eye tracking data confirmed the known split attention problem with overview visualizations [8]. Participants focused a large amount of time on the overview (gaze on overview 20%, SD = 15; gaze on detail 68%, SD = 18). In the *FrameFold* condition the folded areas were focused during 4% of the task completion time (gaze on folds 4%, SD = 3; gaze on focus region 85%, SD = 7). Three participants reported that they did not make use of the overview visualization in the *Overview* condition. This was also supported by the collected eye tracking data. The other participants shifted their attention in average every 3.1 seconds (SD = 2) to the overview area. In the *FrameFold* condition participants shifted the attention only every 6.8 seconds (SD = 4.2) to the folded areas. The gaze shifts in the *FrameFold* condition therefore correspond to only 45% the shifts in the *Overview* condition.

4.3.2 Reconstruction Task

The mean distance in pixels between the original object locations and the participants' recalled locations for the *FrameFold* condition (M = 136 pixels, SD = 98) was lower than for the *Overview* condition (M = 154 pixels, SD = 114; see Figure 3, right). The mean error for the *FrameFold* therefore corresponds to only 88% the error for the *Overview* interface. However, the Wilcoxon signed rank test showed that this difference is not statistically significant (p > .05).
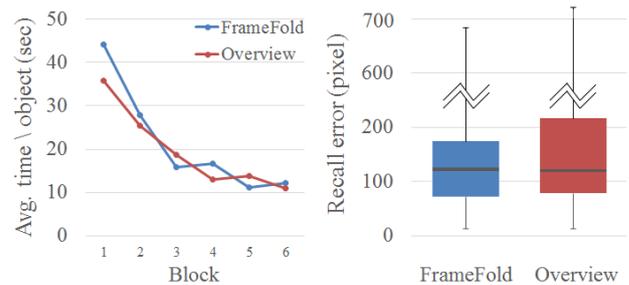

Figure 3. (left) Navigation task: Average navigation time per object for each block; (right) Reconstruction task: Recall error for interface conditions

4.4 Discussion

The task completion time was not significantly different. Therefore also the time the participants had to memorize the object locations did not significantly differ between the interfaces. Although we also found no significant difference for the recall accuracy the results show that the distortion-based interface did not impair users' ability to memorize object locations compared to the undistorted interface.

Whereas we found no significant difference for task completion time and recall accuracy, the attention patterns between the interfaces strongly differed. In the *Overview* condition the participants made excessive use of the overview window to keep orientation. Furthermore they shifted attention between the overview and the detailed view very often. Previous research [8] showed that the separate screen areas can induce spilt attention which we were also able to observe in our experiment. A typical pattern was, that after the participant found an object, he or she had a look at the overview window to encode the object location in relation to the frame of reference. The participants reported that the necessity of frequent gaze shift was annoying because they had to integrate the overview and the detailed view in their heads.

For *FrameFold* participants reported that they did not make use of the folds intentionally. However, participants reported that they perceived the folds in their periphery. This is also supported by the eye tracking data. In the following we discuss the attention patterns of the participants for *FrameFold* in the light of the design guidelines for the structure of the distortion, visual clues and transitions:

(G1) *Scaled-only focus region: FrameFold* distorts the interface only at the border of the viewport while leaving the focus region undistorted. The structure of the distortion allowed the participants to focus most of their time on the easier to interpret undistorted area while using the distorted context only if needed. The data show that participants spend attention on the focus region 85% of the time and that they shifted their gaze only 45% the gaze shift for the *Overview* condition. This indicates that although the participants did not use the folds intentionally they perceived them unconsciously. However, participants did not face a perceived split attention.

(G2) *Real-world metaphors:* Fold-distortions provide a visual clue about the intervening distances [7]. However, in our experiment participants reported, that although they understood the real-world metaphor of a folded sheet of paper they were either not able to interpret the distance accurately enough to help them anchor the absolute position of the viewport or the interpretation of the accurate distance takes too much mental effort to be handled during navigation.

(G3) *Smooth transitions*: Previous research showed that multiscale spaces tax the short-term memory because users must integrate the spatial layout of the information in their heads when panning and zooming [1]. This holds true for the *Overview* condition as participants reported losing orientation sometimes when zooming. In contrast for the *FrameFold* condition participants did not mention any difficulties integrating the pre and post zoom states. The continuous and reversible transition between the folded view when zoomed in and the unfolded view when zoomed out as well as the continuous adaption of the fold depth while zooming allowed the user to keep track of the position within the multiscale space. Whereas the real-world metaphor did not allow for an absolute positioning of the viewport it seem to allow the users to keep track of their relative position while navigating the multiscale space.

5. CONCLUSION
In this paper we applied guidelines for the design of comprehensible distortions on a focus+context technique. We presented *FrameFold*, a technique designed taking into account human abilities to perceive, interpret and comprehend distortions. *FrameFold* shows the whole multiscale space within the viewport while still allowing for pan and zoom navigation. Furthermore we reported the results of a first experiment on how the consideration of the guidelines influence users' ability to build up spatial memory and how the resulting distortion is perceived by the user. In contrast to previous studies which explored fisheye views that do not consider these design guidelines and claimed negative effects of distortions on the ability to memorize object locations, our experiment showed evidence that this does not necessarily hold true for distortion-based interfaces which make use of real-world metaphors, visual clues like shading, smooth transitions and scaled-only focus regions. Our experiment showed that *FrameFold* does not impair users' ability to build up spatial memory **(R1)**. The undistorted focus region and the distortions which were placed at the border of the viewport allowed users to pay attention to the distorted context only when needed **(R2)**. Although the folds did not provide accurate clues about the distance to the border of the multiscale space as a frame of reference they seem to help the users keep track of the current position while navigating the space.

6. REFERENCES
[1] Bederson, B. 2011. The Promise of Zoomable User Interfaces. *Behaviour & Information Technology 30*, 6 (2011), 37–41.

[2] Burigat, S., Chittaro, L., and Parlato, E. 2008. Map, Diagram, and Web Page Navigation on Mobile Devices: the Effectiveness of Zoomable User Interfaces with Overviews. *Proceedings of the International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '08)*, ACM, New York, NY, 147–156.

[3] Carpendale, M.S.T., Cowperthwaite, D.J., and Fracchia, F.D. 1995. 3-Dimensional Pliable Surfaces: For the Effective Presentation of Visual Information. *Proceedings of the Symposium on User Interface and Software Technology (UIST '95)*, ACM, New York, NY, 217–226.

[4] Carpendale, S., Cowperthwaite, D., and Fracchia, D. 1997. Making Distortions Comprehensible. *Proceedings of the Symposium on Visual Languages (VL '97)*, IEEE, 36–45.

[5] Cockburn, A., Karlson, A., and Bederson, B.B. 2008. A Review of Overview+Detail, Zooming, and Focus+Context Interfaces. *ACM Computing Surveys (CSUR) 41*, 1 (2008).

[6] Darken, R.P. and Sibert, J.L. 1996. Wayfinding Strategies and Behaviors in Large Virtual Worlds. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '96),* ACM, New York, NY, 142–149.

[7] Elmqvist, N., Henry, N., Riche, Y., and Fekete, J.-D. 2008. Mélange: Space Folding for Multi-Focus Interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*, ACM, New York, NY, 1333–1342.

[8] Hornbæk, K., Bederson, B., and Plaisant, C. 2002. Navigation Patterns and Usability of Zoomable User Interfaces with and without an Overview. *ACM Transactions on Computer-Human Interaction (TOCHI) 9*, 4 (2002), 362–389.

[9] Ion, A., Chang, B.Y.-L., Haller, M., Hancock, M., and Scott, S.D. 2012. Canyon: Providing Location Awareness of Multiple Moving Objects in a Detail View on Large Displays. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, ACM, New York, NY, 3149–3158.

[10] Mikulecky, K., Hancock, M., Brosz, J., and Carpendale, S. 2011. Exploring Physical Information Cloth on a Multitouch Table. *Proceedings of the International Conference on Interactive Tabletops and Surfaces (ITS '11)*, ACM, New York, NY, 140–149.

[11] O'Hara, K., Sellen, A., and Bentley, R. 1999. Supporting Memory for Spatial Location while Reading from Small Displays. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '99)*, ACM, New York, NY, 220–221.

[12] Ramachandran, V.S. 1988. Perception of Shape from Shading. *Nature 336*, 14 (1988), 163–165.

[13] Sarkar, M. and Brown, M.H. 1994. Graphical Fisheye Views. *Communications of the ACM 37*, 12 (1994), 73–83.

[14] Scarr, J., Cockburn, A., and Gutwin, C. 2012. Supporting and Exploiting Spatial Memory in User Interfaces. *Foundations and Trends® in Human–Computer Interaction 6*, 1 (2012), 1–84.

[15] Skopik, A. and Gutwin, C. 2005. Improving Revisitation in Fisheye Views with Visit Wear. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05)*, ACM, New York, NY, 771–780.