Keeping Things in Context: A Comparative Evaluation of Focus Plus Context Screens, Overviews, and Zooming

Patrick Baudisch, Nathaniel Good, Victoria Bellotti, and Pamela Schraedley
Information Sciences and Technologies Lab/Computer Science Lab
Xerox Palo Alto Research Center
Palo Alto, CA 94304, U.S.A.
+1 (650) 812 4656
{baudisch, ngood, bellotti, schraedl}@parc.xerox.com

ABSTRACT

Users working with documents that are too large and detailed to fit on the user's screen (e.g. chip designs) have the choice between zooming or applying appropriate visualization techniques. In this paper, we present a comparison of three such techniques. The first, focus plus context screens, are wall-size low-resolution displays with an embedded high-resolution display region. This technique is compared with overview plus detail and zooming/panning. We interviewed fourteen visual surveillance and design professionals from different areas (graphic design, chip design, air traffic control, etc.) in order to create a representative sample of tasks to be used in two experimental comparison studies. In the first experiment, subjects using focus plus context screens to extract information from large static documents completed the two experimental tasks on average 21% and 36% faster than when they used the other interfaces. In the second experiment, focus plus context screens allowed subjects to reduce their error rate in a driving simulation to less than one third of the error rate of the competing overview plus detail setup.

Keywords

Focus plus context screens, overview plus detail, zooming, multi scale documents, peripheral vision, experimental evaluation.

INTRODUCTION

Many professional computer users today work with visual documents that are too large and detailed to fit on their screen. No computer display exists today that could, for example, simultaneously display a semiconductor wafer and an individual conductive path on it, because the size ratio between the wafer and the path (five orders of magnitude) would require too many pixels. Furnas and Bederson [12] refer to these kinds of documents as "multiscale" documents.

Multiscale documents are used by professionals in a variety of fields including architecture, graphic design, geographic

information systems, or detailed illustrations (micro-macro readings [20, 5]). Besides static documents (that remain unchanged unless *the user* modifies them), there are also dynamic information *streams* that are too large and detailed to fit on computer displays. For example, a 360° surround view created by stitching the views of multiple cameras will exceed the display capabilities of a typical screen by far.

Figure 1: The f+c interface. The iconic illustration at the bottom left shows where the focus screen is located. The callout shows the different resolutions of focus and context area. The shown document is a map used in the *map task*.

Whenever a document's complexity exceeds what the user's displays can show at a time, users are either required to access information sequentially by manually navigating through the document or to use appropriate visualization techniques. In the related work section, we will discuss sev-

eral solutions that attempt to deal with such documents by better allocating limited display resources to detail and contextual information, most prominently *overview plus detail*.

As an alternative approach, *focus plus context screens* were introduced in [2]. Focus plus context screens are wall-size low-res displays with an embedded high-res display region. Figure 1 shows the display prototype that is currently in use in one of our offices. The frame holds a piece of foam core that serves as a projection surface (4x3 feet). A projector that is located behind the user's back projects onto it, turning the foam core into a large low-resolution display, the *context display*. The foam core has a hole in the center behind which an LCD monitor is located, the high-resolution *focus display*. Customized software is used to display image content across both display regions, such that the scaling of the image is preserved, while its resolution varies across the two display regions. Content scrolled into the focus region is viewed in higher detail, making the focus display behave like a magic high-resolution lens [4].

In this paper, we present the results of a field study, in which we interviewed fourteen diverse visual surveillance and design professionals who all work with multiscale documents or views. This study provided us with a basis for picking appropriate tasks for two laboratory studies to compare user performance with three interfaces for dealing with multiscale documents and views. We found subjects perform faster and more reliably when using focus plus context screens than when using overview plus detail or zooming/panning. We conclude with a brief summary of the results and an overview of future work.

Related work

A substantial amount of research has been done to alleviate the challenge of navigation in multiscale documents. The most basic approach is to use zooming and panning to display required information *sequentially*, e.g., Pad++ [3]. In order to allow users to navigate rapidly between zoom/pan positions, many professional applications allow users to bookmark and recall zoom/pan positions as "views." Another approach is to allow users to open multiple windows that each display a different view of the same document.

Overview plus detail (o+d) visualization is a specialized multi-window arrangement [18,1,14]. In o+d interfaces, one window is the overview, which always displays the entire document. The other window, the detail view, shows a close-up of a portion of the document. The overview typically holds a visual marker highlighting the position of the detail view within the overview. This marker helps to reduce the time required for re-orientating when switching from detail to overview. O+d interfaces have been used for a variety of application areas including medical images [17] and assisting programmers in editing and analyzing code [8]. Shneiderman suggests that the basic two-window o+d visualization is useful for zoom factors (magnification ratio between overview and detail view) of 5-30, after which a cascade of overviews would be better [19].

Focus plus context visualization techniques, such as fisheye views [5,11] or hyperbolic browsers [15] allow users to view selected document portions in additional detail without requiring a second window. However, these techniques introduce distortion, which interferes with any task that requires precise judgments about scale, distance, direction, or alignment. Since many of the tasks in our field study turned out to rely heavily on the accurate representation of distances and scales, we did not include fisheyes in the subsequent lab study.

Focus plus context screens (f+c screens) [2] are the latest work in the field of navigation techniques for multiscale documents. They are similar to fisheye views in that they display content in a single view. However, they do not suffer from the limitations of fisheye views because they avoid distortion. We consider f+c screens as a visualization technique rather than as a display technology. F+c screens are different from wall-size high-resolution displays, in that only a relatively small portion is actually high-resolution, which allows them to be built using less expensive off-the-shelf components. While large resolution displays are desirable, the currently available projection array-based systems [10, 13] are costly, space-intensive, or both. Despite the current progress in combining LCD units into larger displays, such as the 9-mega-pixel display by IBM[1], these displays have not yet reached the sizes and resolutions of hi-res wall-size displays. In the basic conception of using mixed display resolutions, f+c screens have some similarities to variable resolution displays [16], which preserve bandwidth and processing time by reducing the level of detail of visual content towards the viewer's visual periphery, as well as to Feiner's hybrid display consisting of a head-mounted display and a CRT monitor [9].

So far, f+c screens have been demonstrated to be applicable to selected application areas, such as image viewing and editing, 3D gaming, and video conferencing [2]. Since f+c screens display two levels of detail at the same time, we predict that they will help users save time on zoom interactions. This paper presents experimentally derived evidence to support that claim.

1. INTERVIEWS: MULTISCALE DOCUMENT USERS

In order to define realistic tasks for our lab study, we conducted a preliminary field study. We interviewed 14 subjects from different fields: 9 working with static multiscale document and 5 working with views of dynamic content (Table 1). The work of our interviewees covers all five classes of activity in Plaisant et al.'s taxonomy of zooming behavior [18]: *image generation* (e.g., graphic design), *open-ended exploration* (e.g., submarine remote operated vehicle ROV operation), *diagnosis* (e.g., mechanical engineering), *navigation* (e.g., gamers, ROV), and *monitoring* (e.g., air traffic control).

We interviewed all but two people in their usual work setting. We used an informal interview procedure covering the interviewees' experience, the applications they use and their tasks. In all but three cases, we obtained demonstrations of actual work or play on the computer. We also asked about specific problems with existing technology and suggestions for improvement.

Table 1 presents interviewees grouped according to whether they work with static multiscale documents (nearer the top) or dynamic content streams (below), then ordered from top to bottom by the complexity of the document or view, complexity being the size ratio between the document itself and the smallest relevant detail inside the document. As shown by the right-most column we covered a substantial range, ranging from 800:1 to 240,000:1. Using a screen with a resolution of 1024 pixels, a document with 800:1 ratio fits on the screen at once. The 240,000:1 ratio, however, would require users to use a more than 200-times magnification to see the details.

We now summarize our interview findings and the resulting rationale for our experimental tasks. We begin with static document work.

Working with static multiscale documents
Web Design: Our web designer uses zooming capabilities only when using Adobe Photoshop™ to process illustrations and see details in them.

Graphic Design: The graphic designer requires zooming on a frequent basis for large hi-res print products, such as posters. He uses judgment to position, size, and align parts of his product, followed by zooming out to review the overall result.

Mechanical Engineering: The robotics engineer, working with CAD tools, uses zooming to select very small details, such as the individual elements of the boundaries of two closely adjacent components.

Architecture: Like the robotics engineer, our architect mainly uses zooming to simplify selection. Alignment is supported by her CAD tool with a wide range of snapping functions. She is skilled enough at rapid zooming to only require one view even though she has two monitors.

Working with Maps: The GIS specialist, working on land reclamation, works with satellite photographs. She analyzes both archeological and geographic information, which requires frequent zooming to see different levels of detail on a scale from half a meter to entire countries. The zooming is computationally expensive and slow, but this capability is crucial to her work.

Photogrammetry: The first photogrammetry expert we talked to did not require zooming. The second merges maps and verifies compatibility and correctness. She is the only professional user we talked to who uses overview plus detail. In her software, the marker representing the detail view in the overview leaves traces when moved around, so she is actually using it to keep track of which regions she has already covered rather than to navigate.

Chip Design: Our two chip designers work with the deepest documents we have seen. A wafer is 40,000 times the size of its conductor paths and 240,000 wider than one unit on the grid scale on which these paths are created. Their work involves rapid and extreme zooming and panning on a frequent basis to inspect details, place micro-components, make correct connections from one component to another, and to review results.

Deriving tasks for the experimental comparison
Based on our observations we came to a list of tasks that involved a substantial amount of zooming. We excluded tasks that would have required advanced skills, such as image retouching. From the remaining list, we picked one task that involved tracing conductive paths on a circuit board (inspired by tasks in chip design) and one that required comparing distances on a map (inspired by tasks in architecture). We will describe them in detail in the experiment sections.

Working with dynamic views
Air Traffic Control (ATC): Our informant here is an expert in ATC display technology and understands the task of ATC in sufficient detail for our requirements. We learned that ATC workers typically monitor a static sector, guiding traffic and anticipating and accepting or handing over traffic from and to adjacent sectors. Their 2000x2000 pixel spe

	Subject's task	Document/view	Smallest object	Smallest detail	Ratio
Static documents	Web designer	Page: 800 pixel	Text: 10 pixel	Table detail: 1 pixel	800
	Mechanical engineer	Polybot segment: 5cm	Pin: 1mm	Clearance: 0.03mm	2,000
	Graphic designer	Poster: 1m	Text: 1cm	Align: 0.5mm	2,000
	Architect in remodeling	Building: 50m	Wall: 10cm	Accuracy: 1cm	5,000
	Photogrammetry (2)	Highway 2 miles	Curb: 6 inches	Accuracy: 1 inch	100,000
	Geographic info. system	County: 80km	Trees: 5m	Land boundaries: 0.5m	160,000
	Chip designers (2)	Wafer: 12cm	Conductive path: 3µm	Grid: 0.5µm	240,000
Dynamic	Air traffic ctrl. tool builder	Zone: 50km	Airplane: 50m	Plane distance in 25m steps	2,000
	Ego shooter gamer	Surrounding: 360°	Distant opponent: 1°	Aiming: 0.1°	3,000
	Submarine ROV op.	Surrounding: 360°	Small fish: 1cm/0.5°	Use arms: 1mm/0.05°	8,000
	Strategy gamers (2)	Map: 30k pixel	Missile: 3 pixel	Aiming: 1 pixel	30,000

Table 1: The subjects of our field interview and the ratio of the document/view size to the size of the small-

cialized display is considered adequate to see their sector and adjacent sectors at the same time in sufficient detail.

Unreal Tournament™: This is a 3D shooting game and our player has two main objectives: to attack other players with various weapons and to monitor his context for further targets and threats, which he must evade or dodge. His field of view is set to 90°, which requires him continuously to pan in order to monitor his surrounding.

Piloting an ROV: The ROV pilot works from a surface vessel to remotely control a submarine ROV (at depths of up to 1300m). High-power video cameras mounted on the front of the ROV support navigation and other cameras monitor hazards on different sides of the vehicle. He works closely with a co-pilot who will manipulate arms on the ROV, and with a scientist conducting research.

Diablo™ and Everquest™: The two fantasy game players fall somewhere between static and dynamic views. Local activity, such as traveling, casting spells, and fighting are real-time activities, but a substantial aspect of the game involves exploration of the large and relatively static game world. Both programs provide map overviews that show a bird's eye view of a larger context or the entire world to support navigation.

Deriving tasks for the experimental comparison

We ruled out ATC as a possible task to simulate due to a lack of requirements for viewing multiple levels of detail. In the other tasks, dynamic information is viewed at different levels of detail and users simultaneously focus on certain elements while keeping track of possible obstacles, threats or goal points in a wider context. Thus we have a combination of an active "focus task" with a "context monitoring task". For the experiment, we simulated this situation in a driving simulation that we will describe in detail in the experiment section.

Based on the results of our field interviews, we now had realistic tasks that would support a fair experimental comparison between different approaches to displaying detailed and contextual graphical information. In our first experiment subjects worked with static documents, and, in the second, they worked with dynamic content.

2. LABORATORY STUDY ON STATIC DOCUMENTS

In our first lab study, we compared three interfaces, a *zoom and pan* interface, an *overview plus detail* interface, and a *focus plus context screen* interface.

Subjects had to complete two static document tasks. The first task required subjects to decide the shortest path between marked locations on a map of downtown London. The second task required them to verify connections on a circuit board. The goal of this study was to determine which interface would allow subjects to complete their task fastest.

Interfaces/apparatus

The zoom and pan interface (z+p) used a single 1600x 1024 pixel window on a SGI 1600SW flat panel monitor. Figure 2 shows a photo of this setup. Documents were displayed using Adobe Photoshop™. All floating windows and window decorations were disabled. Photoshop allowed the subjects to zoom and pan the document using the mouse (magnifying glass and hand tool). We trained subjects on the tools and shortcuts that they would need to perform the task efficiently.

The overview plus detail interface (o+d) displayed documents in two 1024x768 pixel sized windows (one on each of two monitors, part of each remaining unused). Together, they accounted for the same number of pixels as the single window of the z+p interface. The left monitor (SGI 1600SW) was used for the detail view and the right (Xenon XEN 1810E) for the overview (see Figure 3). Again, we used Photoshop to display the document. The overview was provided by Photoshop's built-in overview palette, which we resized to the desired 1024x768 size. In addition to the navigation supported by the z+p interface, users could pan the detail view also by dragging the detail marker in the overview or by clicking in the overview to re-center the detail view over this point.

The focus plus context screen interface (f+c) consisted of an SGI SW1600 focus display combined with a 1024x768 pixel NEC MT 1035 video projector as the context display (see Figure 1). The focus screen was partially covered by the projection screen to show only a 1024x768 pixel window, so that this setup used the same overall number of pixels as the other two interfaces. One context pixel corresponded in size to approximately 26 focus pixels. Documents were displayed using the f+c image viewer [2] consisting of two instances of the image viewer software ACD-see (http://acdsee.com). Users could pan the document as with the Photoshop hand tool by using a trackball. The f+c image viewer did not support zooming.

Tasks

Overviews always show the entire document while contexts show the local neighborhood surrounding the current focus of attention. In order to prevent this characteristic from influencing the results, we made sure that overviews and contexts displayed comparable information. We achieved this by using documents that were exactly 5285x3964 pixels large. The 1024x768 pixel overview of the o+d interface scaled these images down by a factor of 5.16, which corresponded to the magnification ratio of the f+c context. Therefore, the overview and the context displayed the same information at the same level of detail—whatever was readable in the context was also readable on the overview and vice versa.

The board task: verifying connections on a circuit board
Users were presented with a 5285x3964 pixel bitmap image of a board layout (Figure 2). The document was based on the board design of a PDA-like device that we had modified to make it accessible to non-experts. The three experimental tasks and the training task took place on different non-overlapping subsets of the board.

Each of the experimental tasks was to verify a different set of 24 connections on the board. Connections to be verified were grouped in pairs, each pair clearly marked with a large underlying yellow oval. Each of these connections went from one of the components on the board to a 100-pin chip labeled CPU. Component labels, as well as conductive paths were large enough to be read on the context/overview. CPU pins were labeled in small writing directly on the pin, which made this writing legible only at 100% magnification. Subjects verified a connection by tracing it to the CPU and checking whether the text on the pin matched that at the source. Each set of connections contained 18 correct and 4 incorrect connections. The subjects were instructed to concentrate on speed while completing the tasks. Each subject was thoroughly trained on a training set for each interface before starting the actual task.

Figure 2: In the *board task*, subjects verified connections on a circuit board (z+p interface shown).

The map task: finding the closest hotel on a map
Subjects were presented with one map at a time, each showing a 14x10 km patch of London (Figure 1). The subjects' task was to decide which of two hotels (marked with green and blue double circles) were closer by taxi to the conference location (marked with a red cross) and to call out the answer ("green" or "blue") as soon as possible. The task often required subjects to find their way around parks, rivers, highways, etc., which required subjects to understand map details. Subjects were told to pursue a strategy maximizing both speed and accuracy. There were eight maps to be solved per interface. As in the board task, each subject was thoroughly trained on each interface before performing the actual task. Training consisted of six maps.

Subjects & procedure
The subjects were 12 Xerox PARC employees. We used a within-subject experimental design, i.e., each subject carried out both tasks on all three interfaces. In order to avoid sequence effects, task order and interface order were counterbalanced within subjects.

Subjects performed one task on each interface, then the other task on each interface. They received a verbal expla-

nation and training for each interface. Upon completing the real task, each user filled out a questionnaire containing six general questions taken from the Questionnaire for User Interaction Satisfaction (http://lap.umd.edu/QUIS, [7]) and five questions tailored to the experiment ("I found it easy to orient myself", "The information I needed was easy to locate", etc.). These eleven subjective ratings were combined to form a composite index of user-satisfaction with each interface for each task. Each task (run on all 3 interfaces including training) took between 60 and 90 minutes. The tasks were spread out across two separate sessions.

Hypotheses
Our first hypothesis was that subjects would complete each task faster with the f+c interface than with the z+p interface. The second hypothesis was that the same would hold for comparison with the o+d interface. The first hypothesis was based on the observation that the f+c interface can display relevant micro and macro information at the same time, while the z+p interface requires users to manually navigate in order to acquire this information. The second hypothesis was based on the observation that the f+c interface displays both micro and macro information on a single view, while users of the o+d interface have to visually switch between overview and detail view and reorient themselves after every switch. Because of the visual complexity of the board layout, we expected this effect to be even stronger in the board task.

In the third hypothesis, we expected that the higher efficiency in the tasks would also result in higher subjective satisfaction with the f+c interface and that subjects would prefer this interface in the final rankings. We did not hypothesize that the interfaces would differ with respect to accuracy in performing the tasks.

Results
Table 2 summarizes the average amount of time subjects required to complete the map task (all 8 maps) and the board task on each of the three interfaces. Confirming our hypothesis, subjects achieved best task completion times in both tasks when using the f+c interface. In the map task, task completion took 39% longer when subjects used the z+p interface and 27% longer when they used the o+d interface. In the board task, the differences were 50% and 56%, respectively.

	Map task	Board task
z+p (N=12)	453.3 (236.6)	568.3 (187.1)
f+c (N=12)	325.0 (176.4)	378.8 (71.2)
o+d (N=12)	414.0 (199.9)	593.5 (232.0)

Table 2: Average task completion times in seconds (and standard deviations)

A repeated measures ANOVA evaluating the within-subjects effects of display type (f+c, o+d, z+p) and task (map, board) on completion time found a highly significant main effect of display, $F_{(2,22)}=19.78$, $p<.001$, but no main effect of task, $F_{(1,11)}=2.63$, ns, and no interaction effect,

F(2,22)=1.76, ns. Planned contrasts of display type on completion time revealed that the f+c interface yielded faster completion times than both the o+d (F(1,11)=19.54, p<.001) and z+p (F(1,11)=36.52, p<.001) interfaces. A repeated measures ANOVA evaluating these effects on task accuracy yielded only a main effect of task, F(1,11)=3547, p<.001, but no main effect of display, F(2,22)=2.23, ns, and no interaction effect, F(2,22)=.60, ns. The unusually large F value for task indicates that many more errors were made on the map task than the board task, but this effect is not dependent on different display interfaces.

For both tasks, the majority of subjects (9/12 in the map task and 7/12 in the board task) preferred the f+c interface, which confirms our third hypothesis. A repeated-measures ANOVA examining the within-subjects factors of display and task on the composite user-satisfaction measure yielded a significant main effect of display, F(2,22)=9.30, p<.001. The main effect for task and the interaction were not significant, F(1,11)=0.011 and F(2,22)=0.97 respectively. Planned contrasts indicated that users were more satisfied with the f+c interface than both the o+d, F(1,11)=8.98, p<.05, and the z+p interface, F(1,11)=23.78, p<.001.

Discussion
From our observations, the switching effort between the two views seemed to be the major factor influencing the f+c interface performance advantage over o+d. In the map task, where subjects had to acquire a mental model of the map in order to solve that task, we observed one subject visually switching sixteen times between overview and detail view. When using the f+c interface, this switching seemed much more fluid; the fact that the display areas had different resolutions did not seem to have any effect on subjects' visual navigation. Another factor may have been that subjects found it easier to orient themselves and to locate required information when using an f+c interface compared to the o+d interface, F(1,11)=10.91, p<.01, and the z+p interface, F(1,11)=30.81, p<.001, which we attribute to the consistent scaling across the f+c interface—overview and detail view of the o+d interface display content using two different scales.

However, the experiment revealed two limitations of the f+c prototype used in the experiment. All four users who preferred o+d to f+c for the board task at least partially blamed it on the blurriness of the projection, which made it difficult to trace the conductive paths that were only 1 pixel wide on the context screen. Another problem was shadow casting in the board task. Some subjects, who stood up in front of the projector to physically trace conductive paths using their fingers, found that they cast shadows on the context display, causing them to lose track. Better projection technology and positioning is clearly desirable to avoid these issues.

How do we interpret the performance of o+d in comparison to the z+p interface? Subjects used the overview to rapidly proceed to the next connection to be traced and to monitor progress. This allowed them to keep the detail view zoomed to 100% magnification and to use it exclusively for tracing. When using z+p, some subjects emulated this strategy by zooming out whenever they completed a connection. Although this took extra time, z+p users had a bigger detail view, which allowed them to get further per pan interaction and to complete more subtasks without panning at all. The remaining subjects used the zoomed out view only at the very beginning to *memorize all locations* and to plan the order in which they wanted to complete the task. This allowed them to stay at the most detailed level for the remainder of the task, without any further zooming interactions. This might be why subjects actually performed 4% faster in the board task when they used the z+p interface compared to the o+d interface. However, in the map task, the creation of the mental model of the map required subjects to zoom more often, which allowed the o+d interface to achieve 8% better results than the z+p interface.

3. LABORATORY STUDY ON DYNAMIC VIEWS
In the second lab study, we investigated whether focus plus context screens provide an advantage over overview plus detail solutions when monitoring a dynamic scene.

The car task: avoiding rocks and nails on the street
In the *car task*, subjects interacted with a simple driving simulation implemented using Macromedia Flash™ (Figure 3). The 3D scene consisted of a car on a two-lane street leading through a canyon. The subjects' task was to switch lanes with the car, in order to avoid obstacles continuously appearing on either of the two lanes. Subjects could do so by pressing the left and the right arrow key on their keyboards.

There were two types of obstacles, namely falling rocks and patches of nails on the street. Nails on the street appeared at the horizon and were visible for approximately one second before they reached the car. Rocks came down from the upper rim of the display and were visible for 0.6 seconds before they hit the roof of the car. Collisions were underlined with visual and audible feedback. Nails on the street had the appearance of a noisy gray texture and were not visible on the overview/context, because their texture averaged out when scaled and then exactly resembled the color of the street. Rocks were only visible on the context/overview and measured roughly 200 pixels in diameter. The test sequences were each 120 seconds long and contained 100 fields of nails and 30 rocks in a pseudo random sequence. Subjects were told to concentrate primarily on avoiding the rocks because they would cause five times more damage than the nails.

Interfaces
There were two interfaces. The hardware setup of the o+d and the f+c interface corresponded to their counterparts in the first lab study. The only interactivity was the lane switching, so the o+d interface did not support panning of the detail view or any tight coupling. On the f+c and the o+d interface, rocks were only visible on the overview/ context; nails were only visible on the detail view/focus screen.

Since the speed of approaching nails and rocks would not leave time for zooming, we did not include the z+p interface in this experiment.

Figure 3: The o+d interface during the car task. The left screen shows the detail view through the car's windshield.

Subjects and procedure

Subjects were a subset of eight people from the subject group of the first experiment; the other four subjects were unavailable. The procedure was the same. For each interface, subjects received training (a 2-minute training pass), carried out the actual task, and filled in the same questionnaire as before. Again, a final ranking of the experimental interfaces completed the experiment.

Hypothesis

Our hypothesis was that subjects in the car task would produce lower error rates when using the f+c interface compared to the o+d interface. Using the f+c interface, subjects would be able to extract information about the presence and position of falling rocks using peripheral vision, i.e. without taking focus away from the nails on the streets. When using the o+d interface, subjects would have to visually switch between both views. Again, we expected that the higher accuracy in the car task would also result in higher subjective satisfaction with the f+c interface.

Results

Table 3 summarizes the average number of collisions subjects caused when using the individual interfaces.

	o+d	f+c
Run-over nails	20.9 (10.3)	9.5 (4.1)
Rocks hit	6.6 (3.1)	1.9 (1.9)

Table 3: Mean number of collisions subjects caused in the car task (and standard deviation).

Subjects achieved by far a lower number of collisions when using the f+c interface. Subjects using the o+d interface ran into more than twice the number of nails. The number of rock hits, which subjects were told was the more important type of collision, was 254% higher using the o+d interface compared to the f+c interface. A repeated measures ANOVA examining the within-subjects factors of display (f+c and o+d) and collision type (rock and nail) on collision rate yielded a main effect of collision type, $F_{(1,7)}=843.05$, $p<.001$, whereby subjects hit more nails regardless of display type, but also a main effect of display, $F_{(1,7)}=19.71$, $p<.01$, whereby subjects using the f+c display had fewer collisions overall.

The user's preference ranking mirrors the strong results from the previous study in that all eight subjects preferred the f+c interface to the o+d interface. In addition, subjects reported higher satisfaction on the composite measure for the f+c display compared to the o+d display for the game task, $F_{(1,7)}=23.65$, $p<.01$.

Discussion

By monitoring the subjects' eyes using a video camera, we were able to confirm that subjects' excellent performance with f+c was indeed due to their use of peripheral vision. All 8 subjects continuously stared at the nails appearing in the focus screen without ever looking up. They all confirmed perceiving the rocks and the nails simultaneously. This simultaneous perception allowed subjects to handle the situation with the concurrent appearance of rocks and nails successfully, while they often failed when using the o+d interface.

When using o+d, subjects showed a more diverse spectrum of behaviors. Five out of eight subjects did as we had expected and moved their eyes rapidly between the two screens. Four of them primarily monitored nails on the focus screen and switched to the context screen whenever the nails gave them a break. Only one subject did it the other way around, which allowed him to achieve better scores with the rocks, at the expense of getting the second worst collision rate with nails.

What surprised us were the remaining three subjects. They showed the same staring behavior we observed with the f+c interface and upon request they confirmed that they used peripheral vision to keep track of the rocks. On average, these three subjects collided with only 5 rocks, while subjects who visually switched between views collided with 7.6 rocks. However, all three subjects still did much better when using the f+c interface, where they collided with only 2.3 rocks in average. The reason may be as follows. First, the overview of the o+d interface occupied a smaller part of the subject's field of vision, which made it harder to read. Second, the overview had a much worse reference system. While rocks on the f+c context appeared either left of or right of the subject, all falling rocks on the overview were right of the subject—because the entire detail view was located right of the subject. The fact that the three subjects could still make sense of the situation on the overview was due to the image of the car on the overview, which served as a reference point for deciding on which side the rock was about to come down on.

CONCLUSIONS

We have presented two experiments evaluating focus plus context screens in comparison to overview plus detail and

zooming/panning. Tasks were picked based on the results of a field study, also briefly presented in this paper.

The first lab experiment delivered evidence about the performance of focus plus context screens in the context of static multiscale documents. While the performance of subjects using overview plus detail was only marginally above or even below the performance with a zooming/panning interface, focus plus context screens led to significant time-saving (21% and 36%) in the two experimental tasks, as well as higher subjective satisfaction.

Interaction with dynamic views was the subject of the second lab experiment. We found that f+c screens allow subjects to successfully overload one work task with an additional monitoring task. Even in situations with conflicting objectives (concurrent appearance of rocks and nails), the simultaneous accessibility of all relevant information through peripheral vision enabled users to keep on top of things and to react appropriately. The two-monitor setup provided much less support for peripheral vision and which led to a 254% higher collision rate. These results suggest that f+c screens enable individuals to carry out combined interaction/monitoring tasks that are typically carried out by teams of at least two users, as is the case in the control of a submarine ROV.

How big performance improvements outside the lab will be and whether they will justify the additional hardware effort and space requirements of focus plus context screens remains to be investigated. In future work, we plan to examine the benefits and usage of f+c screens over a longer term in a real world setting and with more applications.

REFERENCES

1. Baldonado, M.Q.W., Woodruff, A., & Kuchinsky, A. Guidelines for Using Multiple Views in Information Visualization, in *Proc. of AVI '2000,* 110–119

2. Baudisch, P., Good, N., Stewart, P. Focus plus context screens: combining display technology with visualization techniques, in *Proc. of UIST 2001*, 31-40.

3. Bederson, B. B., and Hollan, J.D. Pad++: A zooming graphical interface for exploring alternate interface physics, in *Proc. of UIST' 94*, 17–26. ACM press.

4. Stone, M.C., Fishkin, K., and Bier, E.A. The movable filter as a user interface tool. *Proceedings of CHI '94*, 306–312, ACM Press, New York, 1994.

5. Card, S., Mackinlay, J, and Shneiderman, B. *Readings in Information Visualization, using vision to think*. Morgan Kaufmann, San Francisco CA, 1999.

6. Carpendale, M.S.T., Montagnese, C. A. Framework for Unifying Presentation Space, *Proceedings of UIST 2001*, 61-70. ACM Press, New York, 2001.

7. Chin, J., Diehl, V., and Norman, K. Development of an instrument measuring user satisfaction of the human-computer interface, in *Proc. of CHI '88*, 213-218.

8. Eick, S. G., Steffen, J.L., and Sumner, E.E. (1992) Seesoft—A Tool for Visualizing Line Orientated Software Statistics. *IEEE Transactions on Software Engineering*, 18(11-Nov), 957-968.

9. Feiner, S. and Shamash, A. Hybrid user interfaces: breeding virtually bigger interfaces for physically smaller computers, in *Proc. of UIST' 91*, 9–17.

10. Funkhouser, T. and Li, K. On the wall… large displays. *IEEE Computer Graphics & Applications*, 20(4), July/August 2000.

11. Furnas, G. Generalized fisheye views, in *Proceedings of CHI '86*, ACM Press, 16–23.

12. Furnas, G. W. and Bederson, B. Space-scale diagrams: understanding multiscale interfaces, in *Proceedings of CHI '95*, ACM Press, 234–241

13. Guimbretière, F. Stone, M. and Winograd, T. Fluid Interaction with High-resolution Wall-size Displays, in *Proceedings of UIST 2001*, 21-30, ACM Press.

14. Hornbaek, K., Frokjaer, E. Reading of electronic documents: the usability of linear, fisheye, and overview+detail interfaces, in *Proc. of CHI '01*, 293–300.

15. Lamping, J., Rao R., and Pirolli, P. An F+c Technique Based on Hyperbolic Geometry for visualizing Large Hierarchies, in *Proceedings of CHI'95*, 401–408.

16. Loschky, L. C. and McConkie, G. W. User performance with gaze contingent multiresolutional displays, in *Proc. of Eye Tracking Research & Applications Symposium* 2000 (Palm Beach Gardens, FL), 97–103.

17. North, C., Shneiderman, B., Plaisant, C. User Controlled Overviews of an Image Library: A Case Study of the Visible Human, in *Readings in Information Visualization*, Morgan Kaufmann, 570–578.

18. Plaisant, C., Carr, D., & Shneiderman, B. Image-Browser Taxonomy and Guidelines for Designers, *IEEE Software*, 12(2):21–32, 1995.

19. Shneiderman, B. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Third edition. Reading MA: Addison-Wesley, 1998.

20. Tufte, E.R. *Envisioning Information*. Cheshire, CT: Graphics Press, 1990.