Focus plus Context Techniques for Pico-Projection based Interaction

Jens Maiero, Ernst Kruijff, André Hinkenjann, *Member, IEEE,* and Gheorghita Ghinea, *Member, IEEE*

***Abstract*—In this paper, we report on novel zooming interface methods that deploy a small handheld projector. Using mobile projections to visualize object/environment related information on real objects introduces new aspects for zooming interfaces. Different approaches are investigated that focus on maintaining a level of context while exploring detail in information. Doing so, we propose methods that provide alternative contextual cues within a single projector, as well as the potential of zoom lenses to support a multi-level zooming approach. Furthermore, we look into the correlation between pixel density, distance to target and projection size. Accompanying the techniques, we report on multiple user studies in which we quantified the projection limitations and validated the various interactive visualization approaches. Thereby, we focused on solving issues related to pixel density, brightness and contrast that affect the design of more effective, legible zooming interfaces for handheld projectors.**

***Index Terms*—Focus plus context, spatial augmented reality, zooming interfaces, projection based systems**

I. INTRODUCTION

MOBILE projectors not only support flexible ways of projection, but also quasi ad-hoc interaction with real-life environments. Yet, they remain limited with respect to their projection abilities. To design well-performing interactive visualization applications for pico projectors, projection characteristics thus need to be reflected. In this paper, we approach the design of more effective techniques by looking at the strengths and weaknesses of pico projections and search for new methods for providing and interacting with contextual information. Doing so, we particularly look into enhancing established flashlight metaphors like [1] [2] [3] [4] to explore information spaces, as these metaphors are well suited for pico projectors. In our approaches we combine zooming principles from proxemic interaction [5] with focus plus context visualization principles to extend the interaction possibilities in reflection to projector limitations. The main premise is straightforward: when the projector is moved closer to an object or surface, more details are revealed as the pixel density, brightness and contrast increases, while moving away provides the user with context (overview). This particularity also raises an interesting question: how can we maintain some level of context when looking at details of a dataset? While not all tasks will require context-preserving methods, we assume that in particular the search through more complex datasets will benefit from contextual cues to direct the search. However, designing context-preserving methods for pico projectors is challenging, as the size of projection is constrained by visual aberrations and brightness and contrast limitations.

The development of the reported techniques went through two iterations. In the first step of every iteration, we investigated the general perceptual boundaries of pico projection as a trade-off of projection limitations, and explored the role of physical objects and audio cues (also called audioscapes) to preserve a certain level of context. Informed by the results, in the second step we looked into the potential of zoom lenses to support a more flexible search behavior with the pico projector. As a result, the main contributions of this article are: (1) formalization of visibility and legibility issues of mobile projection, (2) novel focus+context methods that make use of alternative context and (3) a novel multi-level zooming approach for interactive mobile projection.

II. RELATED WORK

The techniques presented in this paper touch upon various related areas we will now report on.

Focus plus context. The combination of contextual and detailed information relates directly to focus plus context (f+c) visualization techniques and display types [6]. Multiple displays have been combined to create an effective yet spatially constrained setup [7], while also supporting more flexible usage while conveying focus and context [2]. As part of these studies, it was shown that a mobile projector (focus, detail) can be moved relative to a static large projection (context) [1], illustrating an important aspect: while the projector is moved closer to a projection surface, the pixel density (PPI), brightness and contrast of the projection on the surface increases, which is important in order to display legible and well visible details. Weigel et al. introduced different levels of detail that were stated to reflect the pixel density [1], however the authors did not report on how graphical and textual details were optimized for visibility and legibility while making full use of potential detail. The studies reported in this article look particularly at the quantification of these relationships to optimize the usage of pico projectors for interactive, f+c driven visualization.

Augmented reality. Our interface approach extends the usage of the flashlight metaphor applied in mobile augmented reality interfaces [8] [9] [10] [11] [12], as density (detail) and contextual information aspects covered by our approach have not been studied before. Projection-based augmented reality requires the rectification of the visualization. In comparison to Li et al. [13] the proposed method uses a ray based approach in combination with a high-resolution tracking system to enable

Jens Maiero, Ernst Kruijff and André Hinkenjann are with the Department of Computer Sciences, BRSU, Sankt Augustin. Jens Maiero and Gheorghita Ghinea are with the Computer Sciences Department, Brunel University, London.

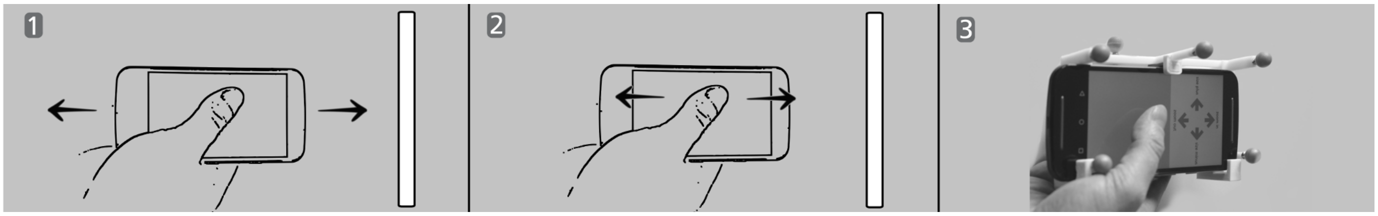

Fig. 1. Multi-level zooming approach. The sketch depicts the pixel density zooming; movements towards/away from the screen increases/decreases the pixel density of the projection (1). Magnifying a region of interest by moving the thumb on the cellphone towards/away from the projection (2). In both proposed approaches the context remains using real world objects, audioscapes or the visualization. (3) illustrates the cellphone on top of the projector.

precision interaction. Furthermore, augmented reality studies often touch upon contextual information cues as this is one of the primary means of conveying information. However, the majority of these studies focus on perceptual issues such as those related to depth perception or occlusion [14]; the actual role of the real world context while exploring specific (sub-) parts of the virtual content is, however, barely studied. While we partly look into context features of physical objects in our study, this is still an area open for further research.

Audio and other non-visual cues. The usage of non-visual cues for providing feedback has found limited application in the field of Augmented Reality [15], including tactile cues [16] and audio [17]. The latter builds upon audification principles that have been explored before to encode visual information [18]. With respect to contextual information provided through auditory cues, 3D audio spaces have been used to provide different levels of information in a spatial setting [19]. Nonetheless, the combination of visual information spaces and audioscapes has not been the focus of much research yet. Using audioscapes offers an additional benefit for focus and context metaphors: as almost all focus and context and zooming interfaces concentrate on single channel feedback [1] [2], using audioscapes context can be preserved even when information is outside the projection area.

Lenses and toolglasses. Finally, one of our interaction approaches deploys a zoom lens metaphor, and as such touches upon the well-explored area of lenses and toolglasses [20]. Lenses have been widely explored as tools to zoom into parts of a dataset and also used to connect specific other functions or visualization methods to adjust the interaction or visualization within the frame of the lens. A thorough overview can be found in [21]. Our implementation differs from the typically used zoom lenses as it is combined with proxemic interaction, allowing a two-step zooming into detailed content, an approach we will describe in detail in the next section. A set of different approaches using projection based spatial interaction combined with lenses have been proposed. The paper lens [22] [23] has been developed in order to improve three-dimensional exploration by moving a white tracked paper in a specific interaction area. A mounted projector projects additional information and different levels of scales onto the sheet. Further approaches focus on lens function, meaning that lens content can be generated separately [24] [25], e.g. time-dependent data. Other issues like different shapes [26] [27] to improve task-specific performance and the categorization of different techniques [28], e.g. explore, select, filter and encode, have been proposed. However, the novelty of our lens is the combination between pixel density, spatial handheld interaction and magnification to explore data in detail.

III. IMPLEMENTATION

As we have shown in the previous section, multiple (handheld) projectors have been used before to convey contextual information while focusing on a sub-set of the information space. However, maintaining a certain level of contextual information while using only a single projector is hard to achieve. If a user zooms into a sub-set of the information space by moving the projector closer to an object of interest, contextual information will get lost, as the size of the projection will reduce. Yet, for more complex search or exploration tasks, it is often necessary to maintain a certain level of context to direct search behavior. Context will provide information necessary to understand spatial relationships, which is hard to achieve when only a sub-set of the information space is visible. In this case, spatial relationships will have to be drawn either from memory or by moving the projector around continuously, reducing search performance effectiveness. Within this paper we explore three approaches to convey context, by using physical objects, auditory cues, or by adding a lens to afford zooming at various levels. Physical objects are unaffected while adjusting the zoom using the projector. As such, users can explore detailed information related to the object without losing the context while zooming into details. Audioscapes look into the potential of non-visual cues, by deploying auditory cues as context. These cues convey location-sensitive information while the user zooms into certain areas. Finally, the third approach affords zoom lens interaction by adding a cellphone to the projector, which could be used to move and adjust a zoom lens within the projection.

A. *System setup*

Our system setup (see Figure 2) deployed a Seeser M2 laser projector (25 ANSI lumen, WVGA resolution). We used a laser projector due to its ability to project sharp, in- focus imaging irrespective of the projection distance. Information was projected on a physical object or a projection canvas. The projection canvas measured 600 x 335 mm, and was made of a wooden panel painted with a matte white paint. Both projector and canvas were tracked using a high-precision ART optical tracking system, tracking at millimeter accuracy. A ray based approach was used to calculate the size of the projection area on the canvas and to rectify the projection.

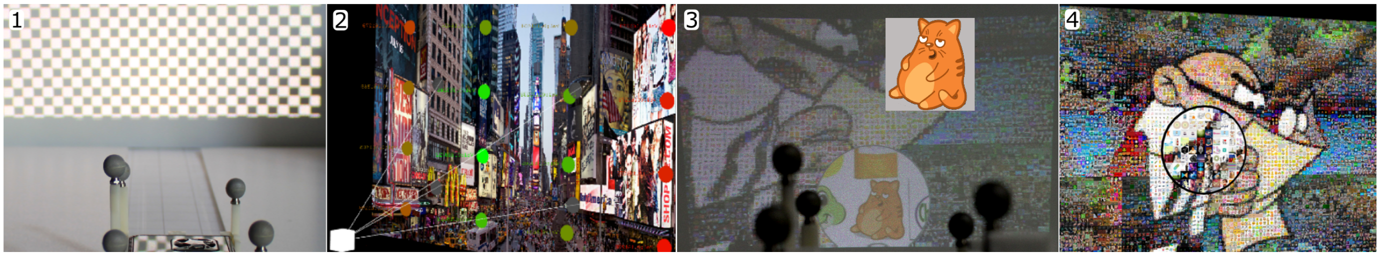

Fig. 2. (1) System overview: the designed user study to estimate visibility and legibility issues. (2) A projection-based image viewer with contextual audioscapes (colored dots for illustrative purposes). (3) System setup of the second experiment depicting the projection area with a located zoomed image and the search image (thumbnail) that was displayed on a notebook next the screen. (4) A mosaic image with zooming lens.

B. Physical reference objects

To study contextual cues, we implemented a zooming interface that closely follows the augmented reality (AR) flashlight metaphor to explore information related to physical objects. In doing so, we deployed standard principles from the spatial augmented reality domain [29], connecting spatially relevant information to real world objects. Based on the inherent image density or by varying the content, different levels of visual detail can be displayed by moving the projector closer or further away from a physical object. To compute, map and correct the projection, objects (with known geometry) were placed relative to a reference system, while radial distortion was removed and the optical center computed. Within our test setup we made use of matte white objects to avoid further aberrations caused by, for example, textured objects.

C. Audioscapes

In our second approach, we explored the potential of spatial audio to provide contextual information cues while zooming into a sub-set of information, as spatial audio is not dependent on projection limitations. Information is encoded in auditory cues that are ordered in a spatially relevant way, similar to [30]. Within our particular implementation, audioscapes refer to the inclusion of small audio regions, called bubbles. These bubbles are dispersed over the full visual scene as depicted by a large image, and thus can also convey contextual information outside the projected area. These bubbles encapsulated distinct sound patterns associated with a certain object in the scene. Proximity and focal point defined the sound intensity of an associated audio file. Looking at the overall image, ambient sound was played (the mixture of all sound bubbles). When a user got closer to the projection canvas, the sound volume of a specific bubble increased, while the surrounding bubbles volume decreased based on proximity. The activation of an audio bubble was based on radius, depending on the distance between the look-at intersection with the screen and the projector's center. The dynamic sound volume was estimated by the inverse fourth power and the normalized distance. Thus, if a user is overlooking the scene (holding the projector further away from the canvas), the volume of all bubbles is about the same as creating a quasi uniform audiospace. In case of the cityscape environment used in our experiment, the audioscape would be similar to the cacophony generally experienced when walking through a large city. Within this particular example, the usefulness of audio bubbles can be well explained: consider vaguely hearing a church bell while observing a sub-part of the scene: the sounds provides a cue that a church may be in the vicinity, after which search behavior can be adjusted.

D. Zoom lens interaction

While moving the projector closer to a projection surface, pixel density, brightness and contrast increases. However, depending on the detail defined by the image resolution, not all details can be displayed once the resolution is higher than the pixel density afforded by the projector. Based on this limitation to handle detail in high-resolution imagery without additional means of magnification, we extended the system with several zoom lens techniques [20]. A cellphone touchscreen (Motorola Moto G2) was used to control the various parameters of the zoom lens, and showed feedback in case users were unsure of the system mode. While the projector itself has some control possibilities, they are very limited, for which reason we extended the basic setup with a cellphone on top of the projector unit (see Figure 1). The cellphone could be comfortably controlled by thumb, without physically obstructing the cooling fan of the projector. The quintessence behind using a zoom lens was to use a ”double zooming” analogy: users could first explore a subset of the scene while still maintaining the overall context of the scene. Once a specific search area or target was thought to be found, users could move the projector closer to the surface, improving the PPI, brightness and contrast, looking at potential details. Hereby, one key aspect should be regarded: it will not suffice to just use a zoom lens to explore details of a scene, as due to projection limitations, not all details can be interpreted when holding the projector further away from the projection canvas. As such, the lens-based interaction is highly affected by the interplay between general exploration of sub-sets from a distance while maintaining context, and moving the projector closer to improve visibility and look closely at details. It is exactly the boundaries between zoom lens usage and moving the projector closer to the canvas that the second stage of experiments focuses on, as reported later in this paper.

Two different types of zoom lenses were implemented. The first type (touch lens [tl]) allowed the adjustment of either the zoom lens size (sliding left/right) and magnification (sliding up/down), or its' placement in the scene hence within the projection screen (sliding all directions). Each of the two modes was accessible by double clicking on the top part of the touchscreen. The second lens type (frustum lens [fl]) made use of the flashlight metaphor, by fixing the zoom lens to the direction the projector points towards the surface. Here, the zoom lens size and magnification could be controlled but, naturally, not its placement within the view frustum, as the zoom lens was simply fixed to the center of the screen.

IV. RESEARCH QUESTIONS

We created two series of experiments. Each consisted of two stages: the first stage targeted the validation of visibility and legibility limitations of the used projection technology, followed by the usability and performance of technique implementations. In the first experiment, we explored the usage of physical objects and audioscapes. The results thereof informed the design of the zoom lenses, which was focused on in the second experiment. Through the experiments, we addressed the following research questions:

R1. Visibility and legibility: how does projection distance relate to perceivable detail in text and graphics?

R2. Context: how well can audio and, to a certain extent, real objects be used to convey context while zooming into details?

R3. Zooming: how do zoom lenses affect search performance while using pico projection techniques?

V. EXPERIMENT 1

In the following, we will describe the experiment setup, procedures and results of both experiments.

A. Apparatus

The first stage of experiments deployed used the introduced pico projector system. The experiment was performed in a darkened environment with controlled lighting (approximately 0.5 LUX at center of the canvas) to guarantee similar conditions between all participants.

In the legibility and visibility stage, a head support was used to fix the user's head so that the eyes would be at center-height of the projection canvas at exactly 850 mm distance. Full screen (overview, context) and maximal focus (detail) was accessible by moving the projector within 750 to 100 mm screen distance. These distances were chosen to comfortably enable interaction within arm-range. The projector was fixed between two bars to allow the movement of the projector forward and backward on a fixed trajectory relative to the projection canvas to create a controlled environment. Furthermore, we made use of a white plastic pot to project graphical content upon. The pot (18 cm diameter) was made of a diffuse white plastic material, which provided a good projection surface without notable projection aberrations. Finally, two loudspeakers provided audio feedback besides the projection canvas.

B. Procedure and design

Visibility and legibility. In the first stage of the experiment, we targeted the quantification of the visual limits of mobile projection based on distance to the projection canvas (hence, proximity), using a simple testbed. Traditionally, visual detail is defined by the visual acuity of the eye (for example using the Snellen test or Landolt rings [31]). However, visibility of graphics and legibility of text is often limited by projection aberrations. Different visual patterns (lines, checkerboard patterns) and fonts at different scales were displayed on a canvas by the pico projector. The study participants were asked to move the projector towards the projection canvas, to the location where the content would be best visible first. The exact location was logged in mm accuracy, and the associated PPI calculated: the size of the detected pattern would help us indicate which level of detail is visually recognizable at a given distance.

The conditions consisted of six types of each category, hence six checkerboards, six patterns and six font sizes, repeated twice. The line and checkerboard pattern scales varied between 0.6, 0.7, 0.8, 1.25, 2.5 and 10 mm and font size 2, 3, 4, 5, 6 and 12 mm respectively, providing a flexible number of transitions. We used randomly generated letters and, for each size, asked the users to spell them out. The size of the patterns and fonts 2 to 5 mm were defined after performing a pre-test, while patterns and letters 1 and 6 are control data. Size 1 is actually not recognizable, while 6 is clearly detectable from all locations.

Context. To explore the general usability of the first two implementations of context-driven techniques, we performed an exploratory experiment. Participants did not have to perform a specific task besides using the techniques to explore the scene at hand.

To address **physical object context**, we created a simple scenario to explore 3D visual contextual information related to a plants' mesofauna, by projecting graphical content on the aforementioned white plastic pot. Content was corrected to correctly wrap around the pot's round surface. Users could explore the life "inside" the pot from different perspectives, introducing three levels of detail. The size of the plant life was correlated with proximity, showing the smallest creatures when the projector was closest to the pot. The three different zones - which faded in at 700, 500 and 300 mm - simply represent three stages in the range the projector can be operated well from close to the human body up to arm's length. The levels of detail thereby followed the simple principle of detail based on proximity: the closer to the pot, the more detailed the information which was presented. It is important to note that the representations changed at the different levels: instead of depending on pixel-wise detail, at 500 and 300 mm additional content was blend in.

To explore the usefulness of **audioscapes**, we created a set of sound bubbles connected to a large-size image of the cityscape of downtown New York (14875x5547 px). The closer the user moved the projector towards the projection canvas, the more detail could be observed, while auditory feedback was provided through the two loudspeakers besides the canvas. In contrast to the plant study, we solely relied on pixel density instead of adding new content. The cityscape contained 16 audio bubbles distributed evenly over the canvas, allocated and clearly connectable to specific visual landmarks in the scene. Within the frame of this experiment, we performed an

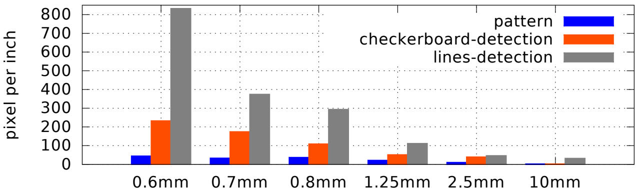

Fig. 3. Comparison between computed PPI for the pattern (gray) and mean PPI when detected by users (checkerboard blue, lines orange - detection rates vary, see Figure 4).

exploratory study to understand the effects of the different kinds of context - especially auditory, but also physical - using the two implemented scenarios (plant and cityscape). Users could freely explore each environment for about 2-3 minutes: there was no specific task besides exploring the various locations and details found in the environment.

After finalizing the experiment, participants were asked to answer 27 questions (using a 5-point Likert scale) that targeted visibility, legibility and information quality issues, the usage of physical object and audioscape context, and general usability issues. The questionnaire included the system usability scale (SUS) 10-point questionnaire to gain insights into the general attitude towards the introduced methods.

C. Results

16 subjects (age 22-47, 4 females, 12 males) participated in experiment 1. The results of the experiment are as follows, ordered along the research questions. The study provided several very useful results, which we will now report on in light of the aforementioned research questions.

R1. Visibility and legibility Human beings can recognize differences between 1 mm from about 3 up to 6 m in real-world environments [31]. Yet, projected content has aberrations that limit visibility and legibility. Figure 4 summarizes the results, while Figure 3 shows the offset between the maximal pattern detail afforded by the projector (as reflected by a lower PPI), and the PPI at which the patterns where recognized by the user (higher PPI, as a product of being close to the projection surface). As expected, participants were able to perceive increased detail closer to the projection canvas by increasing PPI. Figure 4 illustrates that closest to the canvas, line grid patterns smaller than 0.8 mm were only detected by 25% of participants. In contrast, the checkerboard pattern was still visible at around 0.7 mm, affected by a moiré effect at higher PPI. With our particular setup, showing details larger than about 1 mm is not recommended within a projector-canvas distance larger than 280 mm (comparable to 92 PPI) as detection rates drop below 50%.

With regards to text legibility, the results show that text smaller than 3 mm was not legible: it is recommended that the minimum distance between projector and canvas should not be smaller than 330 mm for 3mm size (comparable to 78 PPI). Looking further, Figure 3 shows notable differences between the computed PPI of a pattern and the PPI at which it is visually afforded through projection, considering only those users who detected the pattern (see Figure 4). Clear gaps are noticeable at the higher level details that are only visible at very high PPI, whereas the offset between computed and legible PPI is much lower to almost non-different with lower level details. In conclusion, it is crucial to consider that visual details cannot be shown at their theoretical PPI but need to be adjusted to higher PPI to be fully visible and legible. Thus, performing a calibration step with the used projector before implementing specific visualization techniques is recommended, for which the current setup and procedure can be well used as testbed.

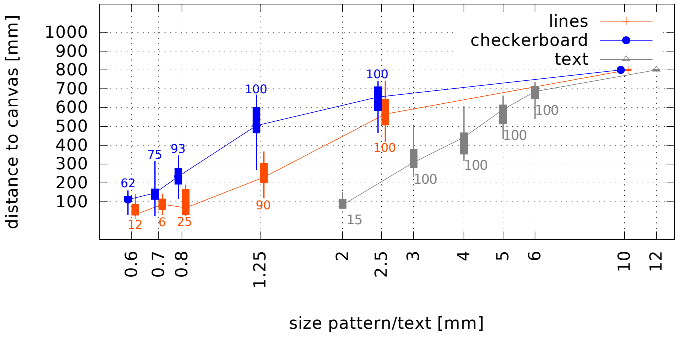

Fig. 4. Results of the visibility test. Numbers indicate detection rate in percentage.

R2. Context With respect to the physical and auditory contextual cues, we analyzed the different implemented with following results. Overall, the SUS analysis revealed that the proxemic-based interface was well received by all users, scoring 84.

Visual quality. Most subjects rated the visual quality of the projector positively. Users were satisfied with the overall sharpness of the projection (avg 3.75/sd 1.90), and noted that the projection quality was reasonable overall (avg 3.00/sd 1.41) and particularly well when the projector was brought closer to the projection surface where the projection notably increased the PPI, brightness and contrast (avg 3.85/sd 1.39). The higher PPI was found to be very important when interpreting detailed information (avg 4.50/sd 1.41). In contrast to the positive rating of brightness and sharpness, many users noted the projection aberrations (avg 1.65/sd 1.56). The deployed laser projector generates a minimal speckle effect and the previously mentioned moiré pattern, in particular with regular line patterns smaller than 1.25 mm. Yet, overall the aberrations only seemed to have slightly affected the subjective view on the system. Finally, the projection size (dependent on projection distance to the canvas) was noted positively (avg 3.56/sd 1.72), which is an important aspect for the flashlight metaphor.

Representation. The information content available at different distances was rated well (avg 4.06/sd 1.49), referring to the switching of plant content and the resolved details in the cityscape image on a pixel-by-pixel basis. The approach of increasing detail in representation (visual and informational level, avg 4.50/sd 1.41), and usefulness of this information for search tasks (avg 4.56/sd 0.99) was rated well, which was on par with the suitability of the level of detail (PPI) based on the distance (avg 3.81/sd 1.62). Hence, our results support the flashlight approach in combination with proximity, as introduced in our implementation.

Physical object context. Users noted that the blending of different levels of information related to real world content was very useful (avg 4.13/sd 1.85) while the real world object was found useful as direct context for the shown content (avg 4.80/sd 0.78). This was supported by the observation of user behavior, as most users explored the plant and its digital contents from different sides. However, more experimentation is required to assess the actual role of physical context, for example in relation to features.

Audioscapes The usefulness of the auditory cues was rated very positively (avg 4.50/sd 1.41), which was likely also affected by the attractiveness of the sound bubbles to enrich the overall experience, an issue we unfortunately did not validate. Subjects noted they could differentiate reasonably well between different sounds (avg 3.68/sd 1.41), even neighboring sounds still being well observable and understandable by most but not all users (avg 3.43/sd 2.23). Similarly, to most but not all users the sound bubbles were well localizable (avg 3.68/sd 2.42). These results support our approach and show that auditory contextual cues can enhance spatial interpretation of the scene.

VI. EXPERIMENT 2

Similar to the first experiment, the second experiment was performed in two stages, looking at visibility and legibility issues and the performance of the implemented technique, respectively. However, in contrast to the first experiment, we assessed task performance using a real search task.

A. Apparatus

The experiments deployed the calibrated environment introduced in the previous experiment. While in experiment 1, the room was fully darkened, in experiment 2 we had constant ambient lighting, to reproduce a more realistic usage environment: the projection canvas center was illuminated by approximately 30 LUX ambient lighting, which resulted in well-visible graphics in the full projection range.

B. Procedure and design

Visibility and legibility We followed up the visibility test of experiment 1 by looking closer at the legibility/readability of a reading text instead of single fonts. We used different font sizes from 3 up to 12 mm, with a step size 1 mm, in reflection to the acquired understanding gained from the first experiment. Each font size was tested once. In the resulting 10 trials, participants were asked to move the projector towards the projection canvas, selecting the distance at which the text was not only visible (legible) but also well readable. No zoom lens was used in this stage of the experiment.

Zooming In the second part of the experiment, the zooming lens implementation was validated, assessing usability and performance in predefined search tasks. We chose to make use of a formalized search task to study performance without directly looking at spatial augmented reality aspects, to better control the task settings and performance. The task was designed so that results can also be ported to the SAR domain. Within this search task, users had to find a specific visual element within the context of a large image. This visual element had the size of a cell in this image, as images were generated as mosaics (see Figure 2). Mosaics were generated with predefined constant cell-sizes (quadratic cells of 32, 64, 96, 128, 160, 192, 224 pixel width). Without zoom lens, each pixel had a projected size of 0,05 mm (equal to 500 PPI or a resolution from 11872x6720 px). Mipmaps were used to reduce aliasing artifacts. As such, cell-sizes were dependent on the level of detail that could be observed with the projector, with and without a zoom lens: details in the smaller cells (128 px and lower) could not observed without lens. The search area was clearly identifiable by the distinct regions in the mosaic, defined by cartoon images. The procedure to structure the search task was as a follows. A laptop was placed right of the canvas that showed instructions on what cell (image) to find in which region in the mosaic image (textual description). Once the user indicated a search pattern (cell) was found, the observer would press enter to create a mark in the search time logging. The projection screen would stay blank as long as the user would read the instructions, showing content only after the user indicated to start the search task, at which logging would start. Cells were partly explored with, and partly without zoom lenses: cell-size from 128 px up were explored without zoom lens, while cells from 128 to 32 px were explored only with the zoom lenses. We thus had one cell-size overlap with the search task without lens. Using the zoom lenses, participants were allowed to magnify content in the range of 1 to 20, and vary the size of the lens between 2 and 20 cm radius, which was kept constant irrelevant from projector - canvas distance: for example, if a user had a 5 cm lens, even if the projector was moved closer to the canvas, it remained the same size. Each cell-size was tested with 3 different search targets, resulting in 36 trials per participant. During interaction, we logged the task completion time, the zoom magnification level, and the location of the projector at which basis we calculated the PPI and the size of the cells. A task was marked as failed when the search target was not found within 2 minutes.

In contrast to the first experiment, we did not use the SUS questionnaire that targeted general usability questions. Instead, we formulated specific questions addressing issues such as projection quality, system ease of use, the appropriateness of the lenses to perform the search task, confidence and concentration. The questionnaire used a 5 point Likert scale, with 5 in positive agreement. During the experiment we also observed participants, specifically looking at user behavior when using lenses.

C. Results

16 subjects (age 23-37, 3 female, 13 male) participated in experiment 2. The results of the experiment are as follows, ordered along the research questions.

R1. Visibility and legibility The analysis of the legibility test reveals that there was a clear cut off (50 %) at 3mm size, below which most participants could not read text anymore

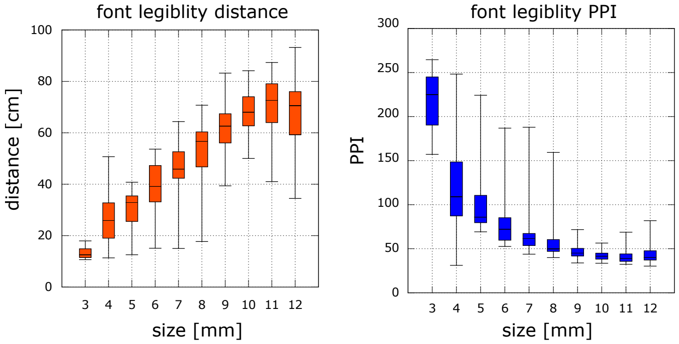

Fig. 5. Legibility/readability test; in this experiment we asked the users to read the text, balancing size of projection with level of detail. The figure depicts the distance and the PPI at legibility level of the 16 participants.

(see Figure 5). Fonts of 7 - 12mm are read at about the same PPI: as a direct relationship between projection distance and PPI, both are interrelated approximately by a quadratic function. At the font size 3-6 mm, participants moved the projector very close to the projection canvas to be able to read the text. Generally, 4-6mm was found to be the minimum size at which text could be well read. In the questionnaires, participants reported that the projector was reasonably well suited to perform this visibility task close to the surface (avg 4.31 / sd 0.79), further away from the canvas the rating lowered, likely because of the brightness and contrast adjustment (avg 3.06 / sd 1.12). Participants were reasonably content with the level of visual aberrations (avg 3.75, sd 1.18).

R3. Zooming Based on 576 trials and the questionnaire outcomes, several interesting results with respect to the usage of the two zoom lenses in comparison to proximity based zooming can be shown. Especially in the smaller cell sizes, the search task was reasonably complex. We noted a clear increase of search time and non-found search targets when the cell size decreased. However, this was also dependent on the differences between search targets themselves: we had larger differences between different search targets in the same cell size (see Figure 7). Search performance time increases with lower cell sizes; however, a clear function is not visible (see Figure 6). This result can likely be explained by the differences in search task patterns. By visually comparing the target cells to their neighbouring cells, higher contrast search targets that were placed in regions that had elicited clear spatial cues (such as high contrast borders) seemed to be found more easily found than other targets. Yet, low contrast search targets and color range issues of the projection did seem to have increased search time, with the extreme case being a target in cell size 224.

A highly interesting finding of the analysis of the cell size (projected) at which the search target was found (see Figure 7) is that participants magnified the cells to a similar size (approximately between 7 and 8 mm), irrespective of the actual size of the cell in the image data, revealing a nice tendency in their behavior. Further informal observations of user behavior by the observer also indicated better performance by participants who developed a search strategy. For example, those participants who seemed to follow distinct visual elements like borders seemed to be faster than other users. However, this kind of behavior requires further methods of logging and analysis not currently possible by our implementation.

	no	fl	tl
		low resolution cells	
ease of use	avg 4.56/sd 0.73	avg 3.94/sd 1.18	avg 3.56/sd 1.03
suitability for search task	avg 4.44/sd 0.81	avg 3.88/sd 1.15	avg 3.69/sd 1.01
precision of control	-	avg 3.63/sd 1.02	avg 3.44/sd 1.03
		high resolution cells	
ease of use	-	avg 3.63/sd 1.09	avg 3.31/sd 1.20
suitability for search task	-	avg 3.56/sd 1.15	avg 3.19/sd 1.28
precision of control	-	avg 3.44/sd 0.96	avg 3.13/sd 1.02
		all resolutions	
learnability of system	avg 4.94/sd 0.25	avg 4.31/sd 0.79	avg 3.75/sd 1.0

Table 1. Results of the questionnaire grouped by lens types and cell resolution

Comparing tasks performed in the first half of the experiment with the second half, most participants improved performance, thus showing positive learning effects. Comparing the overall averages of the different interfaces in the first and second half, users improved performance in the case of no lens (34%), frustum lens (31%) and touch lens (21%). When analyzing the relation between average performance and cell sizes independently from the set of zooming interfaces the performance increased continuously up to a cell size from 96 px. The time for the search task increased from 32 px to 64 px by 23 sec and 64 px to 96 px by 7 sec. From 96 px to 224 px the average performance does not indicate significant changes (+/- 3 sec). Questionnaire data revealed that no lens, as expected, was most easy to learn, while frustum lens was easier to learn than the touch lens. In the lens-only trials, users were generally faster using the frustum lens than the touch lens (45%), which was also supported by the ratings received in the questionnaires (see Table 1). Generally, no lens was preferred by the users and frustum lens scored somewhat higher than touch lens, in the categories ease of use, suitability to use the lens for the search task and even precision of control.

VII. DISCUSSION

Drawing upon the results from the two experiments, we can reflect on the research questions as follows.

R1. Visibility and legibility Overall, the setup used for the visibility studies can be seen as a testbed that can also be replicated for other projector setups. Most flashlight metaphors that have been proposed do not correlate pixel density and proximity. While other proposed systems make use of proximity and therefore different PPI and brightness levels [1] [32], visibility and legibility aspects of mobile handheld projections has not been studied yet. Cao and Balakrishnan defined information spaces in relation to proximity [2], Winkler et. al. introduced a personal (near) and a public space (far) [12], however, Cao and Winkler et al. did not correlate the spaces to pixel density. BaseLase [4] is an interactive focus+context display that makes use of two different projections, a large projection with low resolution to provide the context and a smaller projection with high resolution to provide the focus. This system benefits from the 2 resolutions, but in comparison to our methods BaseLase provides no information about visibility and legibility. In contrast to previous research,

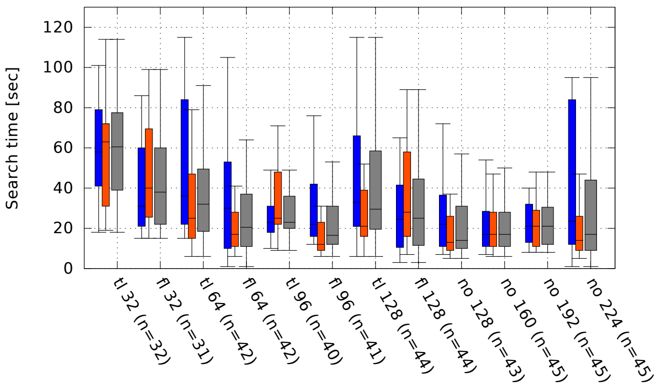

Fig. 6. Search time performance of all participants. Each block represents performance of the first half (blue), the second half (orange) and both together (gray). All trials that last more than 1.5 times of the interquartile range are considered as outlier. The abbreviations refer to the three interaction metaphors: touch lens [tl], frustum lens [fl], no lens [no] and the corresponding size of the patch

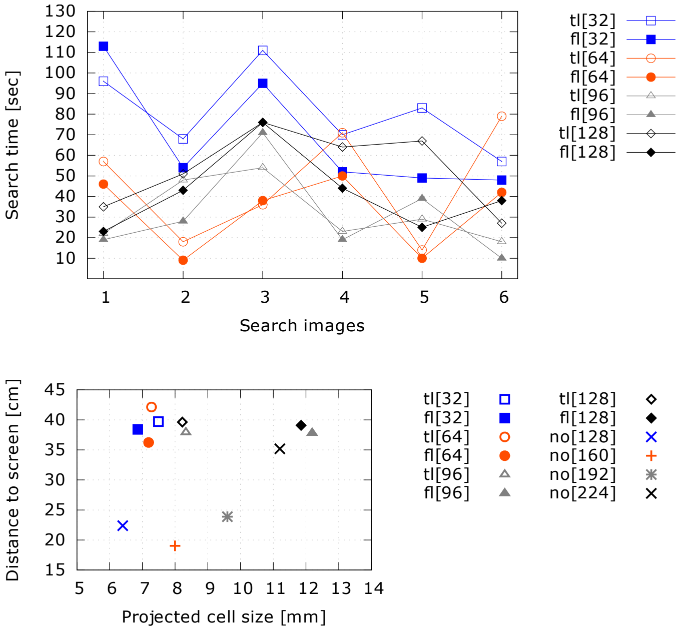

Fig. 7. The graph (top) depicts the average search times, the different scales and zoom interfaces per image. The Figure (below) shows cell size and distance at detection. (tl[x], fl[x], no[x] are the 3 lens interfaces and x is the patch size in pixel)

our results of the performed studies are tied to the deployed projector, but other projectors likely have similar aberrations that can be well verified using similar methods. The results of the two visibility and legibility studies were complementary. Results showed that indeed there is a considerable offset between the theoretical projected PPI and the details that can be conceived by the user. There are limitations in text and graphical visualization affected by projection aberrations that need to be reflected when defining the best level of detail at different proxemic levels. In spite of fonts being identifiable at around 3mm, most participants showed that for reading text, 4-6mm font size is the minimum. Though brightness and contrast likely affected these ratings, we expect that PPI will have been the predominant factor influencing legibility. This quantification can aid in designing interfaces that make use of a similar projector, while the methodology also gives guidance when a different projector is used.

R2. Context As to be expected, users noted physical objects provided a level of context, however, the real effects could not be assessed. It will be highly interesting to more closely study the role of physical objects through formal experimentation. One important aspect would be the interrelationship between information features and physical object features, as the spatial relationships between both will be important to build up a mental model that underlies task performance. It would be worthwhile to contrast physical objects of different levels of feature richness to address their role in performing the task at hand. Research in this area could also drive forward specific AR visualization methods optimized for different feature richness levels. Ni et al. [10] explored in-clinic communication between doctors and patients using mobile projections. The authors provide three different types of context: a wall, a model and the body to project medical content onto it. The performed user study showed that real context (body) supports understanding information and helps to accurate locate the content. In contrast to Ni et al. we extended this approach within our first study and changed the displayed content based on the proximity, which was rated as very useful. Schoening et al. [11] introduced map torchlight a system that augments a paper-based map with additional information like parking lots and further points of interest. In contrast to our study Schoening et al. performed a performance study. In the study they asked the users to find special places on the map. They compared map torchlight with a magic lens approach, where the first performed much faster. This means that augmenting real context could provide not only a better understanding of data, but also could improve performance.

Users also rated the audioscapes very positively. Interestingly, users did not seem to notice that the audio in the interface only communicated proximity and not directional cues. However, since the study was explorative, we only have obtained a general notion on the usability of our approach. While the reception of the techniques was positive, further testing is needed to pin down the actual performance of the techniques. In contrast to Audio stickies [17] we showed that using our approach could additionally provide a valid context and underlines focused areas. Langlotz et al. introduced a similar approach in the area of location based services, however their intention was to provide information about a selected object and not to provide context.

R3. Zooming To cover for offsets between the maximum PPI and details stored in high-resolution images, we showed that zoom lenses - an obvious first choice to tackle visibility issues - can certainly be useful, yet, also come with some restrictions. An important issue to note is that, while we report on zoom lenses in reflection to R3, there are actually some interwoven context issues (R2). Yet, to improve readability, we report on both categories here. The task difficulty level was quite high, which resulted in highly varying performance times, and some learning effects: users performed faster in the second phase of the experiment, which seems to indicate that with further practice, performance can be further improved, which was also noted by some users orally. The projector limitations also showed that solely using the zoom lenses without adjusting proximity with respect to the projection canvas was not possible. Brightness, contrast and PPI was simply too limited to recognize all details when holding the projector further away, even while using higher magnification. The frustum lens performed better than the touch lens, which could be attributed to compatibility with the flashlight approach with which the usage of a pico projector affords well.

Comparing our two proposed zooming methods with other approaches we should differentiate between an interactive lens approaches like [24] and the spatial zooming like [32] which results in higher PPI. Both studies, as mentioned, were explorative so that comparing performance issues is not possible. However, looking deeper into the spatial interaction part of our proposed zooming interface, we can compare our approach with other spatial lenses that uses an stationary projector and a movable lens. In this case our approach clearly reflects the spatial relationship between the lens and the quality of projection. Ni et al. [10] also introduced a spatial lens projecting data onto the human body, while our second interfaces focuses on details in high resolution images we add an second visually lens to enable the users seeing details. To summarize, since zooming is temporal separation our interface supports the metaphor by using an intuitive motion which was also introduced by other flashlight metaphors like [10], [11]. Instead of only using interactive lenses, which is (seamless) spatial separation of the content, our approach combines both.

VIII. CONCLUSION AND FUTURE WORK

Currently, we cannot make any real statements on context interpretation levels in both zoom lens techniques, but this will be an interesting venue for further research: how does motion affect the interpretation of context and, not less important, how is the selected lens size related? While we did not analyze the lens size over time, in further experiments we will do so. There is a dependency between lens size, magnification and available contextual information, as the lens overlaps the context. Furthermore, while participants who searched through content by using some strategy (like following borders) seemed to be performing better, close analysis of this behavior should be studied. In general, further studies should be performed on the actual task chain underlying search tasks, identifying the phases and associated behavior. We assume that large technique improvements can be achieved when more details are known about this chain of action. Another issue requiring a follow up is that magnification directly affected the amount of lens motion of the frustum lens. Small motions resulted in large translations in high magnification mode, and thus could only be controlled with more difficulty. A next implementation will solve this issue, by using a non-linear mapping. Finally, while the formalization of the experiment through cell-in-mosaic search proved useful, experiment procedures can be improved further. In particular the search target features (like contrast) should be better controlled. In addition, while the number of trials of this exploratory experiment was useful for gaining first impressions, further experiments should increase the number of trials to gain a substantial enough basis to assess statistical significance.

In this paper we explored visibility and legibility boundaries of pico-projection based information exploration, and designed user interface techniques to overcome these limitations. Inspired by f+c methods, we looked into context-preserving methods, either by introducing methods to provide contextual information over non-visual feedback, or by allowing the user to zoom into content while simultaneously maintaining a level of context in the visual image. By nature, using a single projector to explore details is somewhat challenging. Moving the projector closer towards a certain area to observe details automatically reduces the amount of context as the projection area gets smaller. To this end, we focused on three important aspects in the accompanying studies: detail, context and zooming. We quantified legibility and visibility boundaries of level of detail and showed, not unexpectedly, that users require larger font sizes to read text than the projector actually affords based on font visibility alone. With respect to contextual cues, the study showed that subjects found the audioscapes useful, while also noting the positive effect of physical context, which can often be establish easily. Furthermore, the zoom lenses were validated positively, yet, also showed they are affected by a learning curve. This shows user performance will benefit more experience with the system, as well as further optimization to the lenses themselves.

REFERENCES

[1] M. Weigel, S. Boring, N. Marquardt, J. Steimle, S. Greenberg, and A. Tang, "From focus to context and back: Combining mobile projectors and stationary displays," in *Proceedings of GRAND Network Centres of Excellence Meeting*, 2013.

[2] X. Cao and R. Balakrishnan, "Interacting with dynamically defined information spaces using a handheld projector and a pen," in *Proceedings of the 19th annual ACM symposium on User interface software and technology - UIST '06*. New York, New York, USA: ACM Press, Oct. 2006, p. 225.

[3] K. D. D. Willis, T. Shiratori, and M. Mahler, "Hideout: Mobile projector interaction with tangible objects and surfaces," in *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*. New York, USA: ACM, 2013, pp. 331–338.

[4] J. Müller, D. Eberle, and C. Schmidt, "Baselase: An interactive focus+context laser floor," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI '15. New York, NY, USA: ACM, 2015, pp. 3869–3878.

[5] N. Marquardt, R. Diaz-Marino, S. Boring, and S. Greenberg, "The proximity toolkit," in *Proceedings of the 24th annual ACM symposium on User interface software and technology - UIST*. New York, USA: ACM Press, Oct. 2011, p. 315.

[6] A. Cockburn, A. Karlson, and B. B. Bederson, "A review of overview+detail, zooming, and focus+context interfaces," *ACM Computing Surveys*, vol. 41, no. 1, pp. 1–31, Dec. 2008.

[7] P. Baudisch, N. Good, and P. Stewart, "Focus plus context screens," in *Proceedings of the 14th annual ACM symposium on User interface software and technology - UIST*. New York, USA: ACM Press, Nov. 2001, p. 31.

[8] S. Kim, J. Chung, A. Oh, C. Schmandt, and I. Kim, "iLight," in *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems - CHI EA*. New York, USA: ACM Press, Apr. 2010, p. 3631.

[9] B. Ridel, P. Reuter, J. Laviole, N. Mellado, N. Couture, and X. Granier, "The Revealing Flashlight," *Journal on Computing and Cultural Heritage*, vol. 7, no. 2, pp. 1–18, Jul. 2014.

[10] T. Ni, A. K. Karlson, and D. Wigdor, “Anatonme: facilitating doctor-patient communication using a projection-based handheld device.” in *CHI*. ACM, 2011, pp. 3333–3342.

[11] J. Schöning, M. Rohs, S. Kratz, M. Löchtefeld, and A. Krüger, “Map torchlight: a mobile augmented reality camera projector unit,” in *Proceedings of The 27th International Conference Extended Abstracts on Human Factors in Computing Systems*, ACM New York, USA. ACM, 2009, pp. 3841–3846.

[12] C. Winkler, J. Seifert, D. Dobbelstein, and E. Rukzio, “Pervasive information through constant personal projection: The ambient mobile pervasive display (amp-d),” in *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, ser. CHI ’14. New York, NY, USA: ACM, 2014, pp. 4117–4126.

[13] Z. Li, K. H. Wong, Y. Gong, and M. Y. Chang, “An effective method for movable projector keystone correction,” *IEEE Transactions on Multimedia*, vol. 13, no. 1, pp. 155–160, Feb 2011.

[14] D. Kalkofen, E. Mendez, and D. Schmalstieg, “Interactive Focus and Context Visualization for Augmented Reality,” in *Proceedings of 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2007.

[15] C. J. Jerome, B. Witmer, and M. Mouloua, “Attention orienting in augmented reality environments: Effects of multimodal cues,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.*, vol. 50, no. 17, October 2006, pp. 2114–2118.

[16] T. T. Ahmaniemi and V. Lantz, “Augmented reality target finding based on tactile cues.” in *ICMI*, J. L. Crowley, Y. Ivanov, C. R. Wren, D. Gatica-Perez, M. Johnston, and R. Stiefelhagen, Eds. ACM, 2009, pp. 335–342.

[17] T. Langlotz, H. Regenbrecht, S. Zollmann, and D. Schmalstieg, “Audio stickies: Visually-guided spatial audio annotations on a mobile augmented reality platform,” in *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*. New York, USA: ACM, 2013, pp. 545–554.

[18] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, “Earcons and icons: Their structure and common design principles,” *Human-Computer Interaction*, vol. 4, no. 1, pp. 11–44, Mar. 1989.

[19] D. N. Zotkin, R. Duraiswami, and L. S. Davis, “Rendering localized spatial audio in a virtual auditory space,” *IEEE Transactions on Multimedia*, vol. 6, no. 4, pp. 553–564, Aug 2004.

[20] E. Bier, M. Stone, K. Pier, B. Buxton, and T. DeRose, “Toolglass and Magic Lenses: The See-Through Interface,” in *Proceedings of SIGGRAPH’93*. ACM Press, 1993, pp. 73–80.

[21] C. Tominski, S. Gladisch, U. Kister, R. Dachselt, and H. Schumann, “A Survey on Interactive Lenses in Visualization,” in *EuroVis State-of-the-Art Reports*. Eurographics Association, 2014, pp. 43–62.

[22] M. Spindler, S. Stellmach, and R. Dachselt, “Paperlens: Advanced magic lens interaction above the tabletop,” in *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ser. ITS ’09. New York, NY, USA: ACM, 2009, pp. 69–76.

[23] M. Spindler, W. Bschel, C. Winkler, and R. Dachselt, “Tangible displays for the masses: spatial interaction with handheld displays by using consumer depth cameras.” *Personal and Ubiquitous Computing*, vol. 18, no. 5, pp. 1213–1225, 2014.

[24] C. Tominski, H. Schumann, G. Andrienko, and N. Andrienko, “Stacking-based visualization of trajectory attribute data,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2565–2574, 2012.

[25] J. Zhao, F. Chevalier, E. Pietriga, and R. Balakrishnan, “Exploratory analysis of time-series with chronolenses.” *IEEE Transactions on Visualization and Computer Graphic*, vol. 17, no. 12, pp. 2422–2431, 2011.

[26] R. Kincaid, “Signallens: Focus+context applied to electronic time series,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 900–907, Nov 2010.

[27] X. Zhao, W. Zeng, X. D. Gu, A. E. Kaufman, W. Xu, and K. Mueller, “Conformal magnifier: A focus+context technique with local shape preservation,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 11, pp. 1928–1941, Nov 2012.

[28] J. S. Yi, Y. a. Kang, J. Stasko, and J. Jacko, “Toward a deeper understanding of the role of interaction in information visualization,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1224–1231, Nov. 2007.

[29] O. Bimber, A. Emmerling, and T. Klemmer, “Embedded entertainment with smart projectors,” *Computer*, vol. 38, no. 1, pp. 48–55, 2005.

[30] G. Eckel, “Immersive audio-augmented environments: the LISTEN project,” in *Proceedings Fifth International Conference on Information Visualisation*. IEEE Comput. Soc, 2001, pp. 571–573.

[31] E. Goldstein, *Sensation and Perception*, 5th ed. Brooks Cole, 2002.

[32] J. Steimle, A. Jordt, and P. Maes, “Flexpad,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI ’13*. New York, New York, USA: ACM Press, Apr. 2013, p. 237.

Jens Maiero is a research associate at the Computer Sciences Department of Bonn-Rhein-Sieg University and at the Institute of Visual Computing, Sankt Augustin. He received the diploma in mathematics from the University of Applied Sciences in Stuttgart, in 2008 and the M.Sc. in computer science from BRSU. Jens is currently working toward a Ph.D. degree in computer science at Brunel University London and BRSU. His research interests include human-computer interaction devices and spatial interaction in smart environments.

Ernst Kruijff Ernst Kruijff is interim professor for computer graphics and interactive environments at the Department of Computer Science of Bonn-Rhein-Sieg University of Applied Sciences, heading the 3D Multisensory interfaces group. His research interest comprises the human-factors driven analysis, design and validation of 3D user interfaces, with a particular focus on multisensory cues. He received his PhD from TU Graz (with honours) and has widely presented his work at the main ACM and IEEE conferences. He is also co-author of the standard reference in his field of research (3D user interfaces: Theory and Practice, 2004).

André Hinkenjann is research professor for computer graphics and interactive environments at the CS department of BRS-U and also adjunct professor at the CS department of the University of New Brunswick in Fredericton, Canada. He is the founding director of the Institute of Visual Computing at BRSU. His research interests include high quality rendering, interactive data exploration and high resolution display systems. André is a regular reviewer for international conferences and journals on the above mentioned topics. He is member of ACM SIGGRAPH, Eurographics and German GI.

Gheorghita Ghinea is a Reader in the Department of Computer Science at Brunel University, UK. He received B.Sc. and B.Sc. (Hons) degrees in Computer Science and Mathematics (in 1993, 1994), respectively, an M.Sc. degree in Computer Science (1996), from the University of the Witwatersrand, Johannesburg, ZA; and a Ph.D. degree in Computer Science from the University of Reading, UK (2000). His research activities lie at the confluence of Computer Science, Media and Psychology. In particular, his work focuses on the area of perceptual multimedia quality and building end-to-end communication systems incorporating user perceptual requirements. He has over 250 publications in leading international conferences and journals; Dr. Ghinea has co-edited two books on Digital Multimedia Perception and Design, and Multiple Sensorial Media Advance and Applications. He is a member of the IEEE and the British Computer Society.