Floor-Projected Guidance Cues for Collaborative Exploration of Spatial Augmented Reality Setups

Susanne Schmidt, Frank Steinicke
Universität Hamburg
Hamburg, Germany
{schmidt, steinicke}@informatik.uni-hamburg.de

Andrew Irlitti, Bruce H. Thomas
University of South Australia
Mawson Lakes, Australia
{andrew.irlitti, bruce.thomas}@unisa.edu.au

ABSTRACT
In this paper we present a floor-based user interface (UI) that allows multiple users to explore a spatial augmented reality (SAR) environment with both monoscopic and stereoscopic projections. Such environments are characterized by a low level of user instrumentation and the capability of providing a shared interaction space for multiple users. However, projector-based systems using stereoscopic display are usually single-user setups, since they can provide the correct perspective for only one tracked person. To address this problem, we developed a set of guidance cues, which are projected onto the floor in order to assist multiple users regarding (i) the interaction with the SAR system, (ii) the identification of regions of interest and ideal viewpoints, and (iii) the collaboration with each other. In a user study with 40 participants all cues were evaluated and a set of feedback elements, which are essential to guarantee an intuitive self-explaining interaction, was identified. The results of the study also indicate that the developed UI guides users to more favorable viewpoints and therefore is able to improve the experience in a multi-user SAR environment.

Author Keywords
Spatial augmented reality, floor based user interfaces

INTRODUCTION
When bringing virtual reality from the labs into our daily life we are faced with a variety of challenges, including technological, ergonomic or social factors. Spatial augmented reality (SAR) has some advantageous attributes to support this transition by providing users with easy access to a virtually enhanced environment, allowing them to benefit from the characteristic features of augmented reality (AR) and virtual reality (VR) without being dependent on heavy and cumbersome to wear optics. The key components of a SAR environment are projectors, which are not in need of specific display screens but instead project the virtual content directly onto real-world objects. Since the display technology is detached from the user, SAR setups get along with a minimum level of user instrumentation. In order to experience 3D projections, light-weight 3D glasses as known from cinemas with attached markers for 6DOF tracking are sufficient. Furthermore, if only monoscopic projections are presented, additional head-worn or hand-held devices are not required at all. An important side effect of this separation between display and user is the capability to provide a shared space for multiple users. Since the augmentations are not displayed on private devices and instead projected into the real environment, a number of people can participate in a collective experience. Another positive aspect of SAR technology is its flexibility regarding the level of virtuality. While most VR/AR applications focus on a selected stage within the mixed reality continuum, the use of projections makes it possible to add any desired amount of virtual content to the real scene and therefore allows to transition seamlessly between different stages.

Due to these advantages of SAR, it may be used for a variety of applications. In particular, experiences in public space or, generally speaking, experiences that involve groups of people could benefit from the seamless, social and connective character of the technology. Figure 1 outlines two possible applications, the presentation of a real exhibit in a museum and the discussion of a physical block model as part of the architectural design process. By using SAR technology, the physical objects can be supplemented with additional virtual content in order to highlight specific details, change the objects' appearance or even show them in their natural context.

Looking closely at these two examples, certain limitations of the technology become apparent, which could negate the initial benefits of SAR. First of all, both scenarios include content that is projected monoscopically onto the surface of physical objects and therefore can be viewed by multiple users without visual distortions. On the other hand, they also include stereoscopically projected content as this allows to present additional information such as the restored version of an artifact or the interior of a building. However, projector-based systems using stereoscopic display are usually single-user setups, since they can provide the correct perspective for only one tracked person. Exceptions are projector systems as presented in [19], which use a high frequency in order to render different views for up to six users. However, such systems are customized, highly complex and usually not affordable for applications that have to cover larger areas and therefore are based on more than one projection unit.
Another limitation of projection-based AR systems is the existence of shadows that might interfere with the projection. Particularly for objects with complex shapes as the dinosaur skeleton, self-shadowing is usually inevitable, unless an excessive number of projectors is used. One option to reduce the visible shadows from the user’s point of view is to correlate his position with the projectors’ frustums. If the user’s head is close to the optical center of the regarding projector, shadows are occluded by the physical object itself and therefore do not disrupt the projection.

Restricting the user’s movement within the scene and directing his gaze to a specific region can also be of practical value from a narrative point of view. As in all immersive VR and AR setups, SAR also faces the challenge to allow users autonomous exploration of the scene and to present pre-configured story elements at the same time. Indeed, users might miss important elements of the story because they are looking in a different direction.

Finally, for a SAR system with different stages, the question arises how to switch between these stages while preserving the simplicity of the system instead of complicating it by introducing unknown interaction techniques.

With regard to these considerations, our goal is to develop a user interface (UI), which emphasizes the positive effects of a SAR environment and reduces possible complications, which may emerge in everyday use of the technology. In detail, the interface is designed to address the following three issues:

1. **User Interaction**
   Allow users to seamlessly transition between different states of the system, without needing to use additional input devices or to learn application-specific interaction methods.

2. **User Guidance**
   Support the storytelling by guiding users to regions of interest and ideal viewpoints.

3. **User Collaboration**
   Extend the system to multiple users by introducing a master-follower concept and corresponding visualizations.

In order to meet the above requirements, we developed a floor based UI, which is adapted for, but not limited to, domains as exhibitions or architectural meetings, in which SAR technology can be used to present different aspects of a physical object. In the next section, we will discuss related projects that focus on floor based UIs, storytelling in immersive environments and collaboration in shared SAR environments. Section 3 presents the proposed floor interface and discusses how particular design elements are supporting the three main goals of the project. Section 4 describes a user study, which evaluates the floor interface in terms of usability as well as social factors. Section 5 concludes the paper and highlights future research directions in this field.

**RELATED WORK**

We sample related work in three separate research areas: (1) Floor based UIs with various interaction and visualization methods, (2) delivering structured stories and communication in immersive environments, and (3) SAR, its design and use in collaborative and view dependent applications.

**Floor Based UIs**

Ubiquitous computing, where computers are seamlessly integrated into the space around us, are a motivation for the design and implementation of interactive floors. Spatially immersive displays, such as the CAVE [13] or the pioneering ideas in *Office of the Future* [26], employ the floor as an extension to the users visual experience. While these experiences utilize static displays, projecting graphical information dynamically across a floor can also be achieved using the combination of a projector and a rotating mirror to direct light [25].

Floor interaction is commonly recorded either through contact sensing [11, 39, 38] or by under-floor camera tracking [14, 3, 8, 32]. Both methods allow for the fine-grained capturing of floor-based touch input either through feet, body, or additional apparatus. Augsten et al. [3] used frustrated total internal reflection (FTIR) with a high camera resolution to capture user input. The rear projected surface presents information such as menus, crosshairs, text, and graphical projections. Users used explicit actions such as stepping, jumping, tapping, or
stomping to interact with the interface, however foot posture could also be used as form of input. The Gravity Space smart room [8] also employs FTIR to track users and furniture, with tracked objects visualized as an imprint across the floor. The work uses algorithms to identify users and poses, demonstrating the ability to identify existing users in the space based on shoe prints stored in a database. Visell and colleagues discuss several interaction techniques for their contact sensing multi-modal floor interface. The floor provides user input for widgets and menus [39], geo-spatial navigation [38], and can also provide vibro-tactile feedback in response to movement [37]. In user interaction evaluations, the authors noted that small interface elements projected too close to the user caused interaction difficulties. Occlusion was also found to be an issue; however top down projections could alleviate this issue. As opposed to direct foot input, Schmidt et al [32] used kickable objects as an interaction instrument for their projected menus. The physical objects could be independently manipulated to operate interface controls such as sliders and buttons, however they could also be manipulated together to create relationships, such as creating molecules in a chemistry application.

For collaborative tasks, the floor can be supportive for its use in both visualizations and interactions. Krough et al. [18] discuss an interactive floor for facilitating Q&A driven by SMS and email. The iFloor combines a ceiling mounted projector and a vision-based solution to track users, using each user’s captured positioning as a magnet for attracting the cursor. A limitation of the system is that it only tracks users in a meter band around the outside of the visualization; the projected area is used only for the transfer of information. Inspired from the iFloor, Grønbæk et al. [14] discuss their system which places four projector-camera pairs beneath the floor, allowing users to interact with projected content across the entire visualization space. Users can interact and participate in collaborative games such as Pong using body interactions. A key design in both examples is the associated communication between co-located participants; using the floor and its associated visualizations as a basis for discussion and instruction.

As opposed to interacting through detecting gestures [3, 8, 14], supplying slippers [12, 20] or requiring additional tools [32], interaction in our system resembles the idea of proxemic interaction, based on the position, identity, movement and orientation of entities in the scene [4].

**Storytelling in Immersive Environments**

There are numerous examples of educational storytelling using mixed reality [1, 10, 24, 21, 31, 29]. The *Ghosts of Sweet Auburn* project [21] used video actors augmented into the environment to deliver a narrative story to users wearing a head worn display. Cyberguide [1] was an early example to the potential of a location-aware device, coupling guidance and communication to enhance a user’s experience. The co-visiting system Lighthouse [10] allowed physical users to experience a museum using a handheld display alongside a user in virtual reality, and another using the web, sharing voice communication, maps and a 3D display. Based on conventional visitors, the authors noted that sharing experiences is achieved through subtle forms of communication, with visitors sharing physical movements and observations as they traversed through the exhibition space. Schmalstieg and Wagner also used a handheld device with fiducial markers to deploy a treasure hunt game inside a museum, linking selected exhibits into an engaging story [31]. It was observed during evaluations that users began collaborating voluntarily, intuitively forming groups to work together before splitting up again.

Miyashita and colleagues similarly used a handheld display but presented animations automatically when users were in proximity to a station. The system received complementary reviews from the public, however it was noted that the system introduced divided attention between the screen and the artwork; numerous users commented on the difficulty to appreciate the artwork without looking through the magic lens.

Ridel et al. [29] used an interaction revealing flashlight, to project expressive visualizations over physical artifacts. Users could use a pointing gesture to direct a cone of light onto the display, presenting additional registered information. The system was deployed in a museum exhibit where the curator commented on the system’s success at remaining hidden and allowing focus to remain on the exhibit.

**Collaboration in Shared SAR Environments**

SAR is inspired by the impact of projecting information into a user’s environment, merging the physical and virtual worlds [7]. Early collaborative examples used projectors to augment the space with additional displays. Augmented Surfaces [28] facilitated the collaborative transfer of digital information between portable computers, table and wall displays. The Digital Desk [40] combined physical and projected elements to achieve a computer-augmented tabletop environment. A user’s interaction with a physical element on the desk (i.e. a numerical entry in a bill) could implicitly update projected elements (i.e. enter the value into a calculator). The *Office of the Future* [26] proposed using all surfaces inside an office space be implemented as displays, allowing the remote transmission of not only users inside the space, but also the space itself. Extending SAR beyond planar projections, Shader lamps [27] uses projected light to modify the visual properties of a physical model; providing flexibility in both visualization and interactive experiences using everyday physical models as displays. Disney has employed this technique inside their theme parks exhibits to immerse visitors inside their magical space [23]. SAR research is predominantly driven from this idea, with numerous examples demonstrating its use in collaborative applications [17, 16, 19, 5].

The *IllumiRoom* [17] maps an environment around a conventional display using depth cameras. The captured space is then augmented with projections to extend the visual display area, focusing on stimulating a user’s peripheral vision. *RoomAlive* [16] extends this concept to encompass an entire room, allowing multiple users to immerse themselves into the display.

We also draw from a number of projects which demonstrate perspective views using SAR [6, 19, 2, 5]. Kulik et al. [19] time-multiplexed the display of six modified projectors to
present 12 distinct images (six users, left-right eye) on a single planar display. The system provided view dependent renderings to each user using shutter glasses. A number of cues were developed to assist in the collaborative navigation and visual acceptance through virtual spaces. Alternatively, the virtual showcase [6] used space-multiplexed beam splitters to allow the presentation of stereoscopic overlays on physical artifacts; four independent views are combined with mirrors and provided around the display. This work is further extended using a curved display, allowing multiple users to seamlessly experience the central display. Benko et al. [5] present a shared face-to-face collaborative environment, using spatially distinct regions of projection. The areas behind each participant provide a suitable space to present view dependent rendering for each user, however the system relies on each user staying opposite one another to maintain correct display registrations. Adcock et al. [2] present a projected environment with multiple complementary views for a single local user viewpoint. Below surface visualizations are projected using view dependent rendering for the local user to simulate cut-outs in the surface, while shadowing in the scene is also presented in relation to a remote user’s positioning. A limitation of this work is that the visual cues only work from the single vantage point. Our work extends on this limitation by providing spatial cues to co-located users to modify their positioning to correct projected views.

DESIGN OF A FLOOR BASED UI

The floor interface was designed to address three design goals: supporting (i) user interaction, (ii) user guidance, and (iii) user collaboration. In the following, detailed information on the design goals and derived interface elements are provided.

User Interaction

The basis for our UI is formed by a scene selection menu, which consists of a number of buttons that are projected onto the floor (see Fig. 2). Each button represents one content-related scene and can be labeled with the according topic. When a user steps into a specific button, a circular progress bar frames the button. When the loading is complete, the application transitions into the selected scene and the button transform into a larger floor area. This area indicates the walking zone and is discussed in greater detail in the following section. While the user moves within the boundaries of the walking area, he can explore the scene autonomously. To exit the current scene and return to the scene selection menu, the user just has to step out the walking area. In order to prevent an accidental leaving, e.g., when the user is moving backwards, the entire floor UI is vibrating as soon as the user is approaching the boundaries.

User Guidance

In order to ensure that users can make the best of a SAR experience, particularly favorable viewpoints are suggested via customized UI elements. At top level, every scene is connected to a pre-defined area that is safe to walk in, both with regard to the storyline and the technical limitations such as self-shadowing. At scene entry it is expanded in stages according to a color gradient, which leads to a two-dimensional floor map showing the quality of different viewpoints, where only areas with a minimum quality level are included. To make sure the user is entering a scene with the ideal viewpoint, the buttons’ layout can be constructed accordingly. For that purpose, the individual position of a button can factor in different aspects of the linked scene. First of all, the scene’s level of virtuality usually has a strong impact on which real and virtual elements are of utmost interest for the user. If no virtual images are overlaid, the objects themselves are brought into focus. Therefore, a close distance to the object could help the user to discover details in terms of shape, material qualities, and surface texture. In contrast, if the object’s context is displayed, a comprehensive view of the entire surrounding might give a better impression of the scene in its entirety. Overall, the distance of the buttons to the physical exhibit can reflect the relevance of specific scene elements and therefore allows the designer of the system to draw the users’ attention to them. This proximity-based approach can be expanded by the direction the user is facing at the beginning and throughout a scene. The goal is to provide unobtrusive cues that suggest regions of interest (ROIs) to the user, rather than forcing him to look in a specific direction. This is achieved by two different interface elements, which are used according to the current state of the system. When the scene selection menu is displayed, footprints inside each button indicate the ideal direction of view for every scene. After a scene was loaded, the footsteps disappear and are replaced by a circular segment.
which represents the current ROIs within the scene. All UI elements for guidance are illustrated in Figure 3.

![Image](image-url)

**Figure 4:** Interface elements to support user collaboration. The optimal distance to a selected master (the user who controls the perspective) is indicated via color-coded circles with an additional arrow.

**User Collaboration**

At this point, a single user can take full advantage of the system, explore the different scenes and transition between them using the floor interface. It is also easy to extend the system to multiple users, if the current scene only contains virtual objects that are independent of the viewpoint (as in the case of the states in Figure 1a, 1b and, as appropriate, also 1d). However, for scenes that contain 3D content as in Figure 1c, the virtual cameras have to be coupled to the pose of an observer’s head in order to convey a realistic sense of perspective when the observer walks through the scene. Therefore, even for multiple users a correct perspective can only be provided for the position of one observer. In order to prevent conflicting user inputs during the exploration of such a scene, one user is dedicated to the master task and therefore controls the perspective for all observers of the scene. This concept was already introduced for user collaboration in projection-based 3D virtual environments without augmentation of physical props and was judged favorably in a user study that focused on subjective measures [33]. Which user is chosen to be the master is decided in the moment of selecting the next state in the main menu of the system. Other users can wear shutter glasses as well, allowing them to perceive the scene stereoscopically. However, in order to get a less distorted perception of virtual 3D models they have to stay close to the master. To inform users about this, we introduce different UI elements as shown in Figure 4. First of all, the master is identified with a gear-wheel around his feet along with the lettering ’master’. Every other user is surrounded by a colored circle, which represents the grade of the user’s current position. In scenes without 3D content the circle is always green. If the current scene contains 3D content and the user is too far from the master to have a good viewpoint, the circle turns reddish. In addition, an arrow appears to show which direction the user has to go in order to improve his perspective. By this means, the master can move freely within the scene while other users are encouraged to follow his movement. By his current position, the master also decides when to leave the current scene and return to the scene selection menu. In the menu, the master task can be passed on to another user as described before. Whether this approach is feasible in a realistic scenario or if an automatic timer to leave the scene as well as a balancing strategy to assign the master task should be implemented, are two of a number of questions we wanted to investigate in a user study.

**USER STUDY**

For the evaluation of our proposed floor interface we simulated an exhibition scenario in a CAVE, with a physical dinosaur skeleton serving as the central exhibit. In order to put emphasis on social aspects of the interface, participants completed the study pairwise, as illustrated in Figure 5. Following a between-subjects design, the interface was compared to a control condition, which was reduced to basic UI elements. The control condition involved plain floor-projected buttons that pulsed to gain the users’ attention and stopped pulsation after one of the users stepped in. In contrast to the developed extended UI, the buttons’ locations were not adapted to the scene content in terms of distance and direction. Also, none of the previous described guidance cues were used in the basic UI.

**Participants**

We invited 40 participants to our study, 26 male and 14 female (aged from 19 to 65, $M = 29.5$) and assigned them to 20 experiment sessions. 32 of the participants were students or staff members of the local computer science department, while 8 participants stated to pursue a non-technical profession. To model a natural situation in a museum, half of the participants already knew their partner while the other half of participant pairs were strangers prior to the beginning of the study. This differentiation will be taken into account during the analysis, in order to identify possible issues of the interface when two strangers have to interact with each other in a shared space. In order to qualify for participation in the experiment, each user had to confirm to be unfamiliar with the CAVE. This prerequisite was used to ensure that participants had no preknowledge over the functionality of the CAVE and its limitations regarding the multi-user capacity.

**Setup**

For conducting the user study we used a 4-sided CAVE with three walls and the floor as projection surfaces [34]. Inside the CAVE, close to the front wall, a replica of a dinosaur skeleton was positioned on a white box. In order to augment the physical object as well as its environment, the CAVE was equipped with five 3D projectors. This involves three Optoma EH320UST projectors, which were mounted on the ceiling at a distance of around 0.8m in front of the walls. For the floor as well as the exhibit, two Optoma GT1080e projectors were used. The latter was mounted on a swiveling arm that was moved close to the exhibit to reduce self-shadowing and to improve pixel density of the projections on the exhibit’s surface (see Figure 5b). Overall, the projections covered a space of $3.15 \times 4.2 \times 2.36m$. Both participants of an experiment session had to wear 3D shutter glasses in order to experience the stereoscopic content. As our intention was to achieve perspective correct projections that appear to surround the real exhibit, an ARTTRACK2 tracking system was installed to measure the users’ head poses.
Design and Procedure

Prior to the study, both participants had to fill in a consent form, including a declaration of the planned video recording. After a small introductory story to stage the experience, participants were instructed to put on the shutter glasses, enter the CAVE and explore the presented exhibition as during a normal museum visit. Apart from this, no specific tasks were given and the used technology was not introduced.

In total, three scenes could be selected as shown in Figure 1 (top). This included a presentation of the dinosaur’s anatomy with 2D highlighting and 3D textual annotations, a 3D projection of the skin around the skeleton, and a stereoscopic 360 degree video that showed the habitat of dinosaurs. All scenes were accompanied by an audio commentary. At any time, participants were free to talk to each other and to move through the CAVE, however, as in usual museums the skeleton must not be touched. The behavior of the participants as well as their conversations were recorded using a video camera. Furthermore, additional data was stored for later analysis, including the distance between users, the distribution of the master and follower roles, and targeted objects. After 8 minutes of free exploration, participants were asked to move into two separate rooms and to fill in some post questionnaires. This included scales regarding usability and subjective communication. In total, one experiment session took around 30 minutes.

Results

A variety of subjective and objective measures was used to evaluate different aspects of the developed UI. In the following, we refer to participants who used the basic UI as the control group and to participants with the extended UI as the test group, respectively.

User Interaction

The usability of the presented UIs was investigated both with the System Usability Scale (SUS) [9] and the AttrakDiff questionnaire [15]. We analyzed the results with five unpaired t-tests at the 5% significance level. Since no significant differences between the test and the control group could be found, we pooled the results of both groups. The average SUS score adds up to $M = 73.250$ ($SD = 14.212$), which can be interpreted as a grade of a B [30]. On a scale of -3 to +3, the pragmatic quality was rated with $M = 0.996$ ($SD = 0.806$), the hedonic quality (identity) with $M = 1.004$ ($SD = 0.907$), the hedonic quality (stimulation) with $M = 1.061$ ($SD = 0.748$) and the attractiveness with $M = 1.600$ ($SD = 0.819$).

In addition to the usability scales, we adopted a questionnaire from [22] to measure the usefulness, accuracy and effectiveness of the floor-projected feedback on a 5-point Likert scale. The results were analyzed using three Mann-Whitney-U tests, since the assumption for normality could not be assumed. We found significant effects at the .05 significance level for the feedback usefulness ($U = 83.000, p = 0.001, r = 0.510$) and the feedback accuracy ($U = 86.000, p = 0.002, r = 0.496$). The effect of the UI type on feedback effectiveness was not significant ($U = 129.500, p = 0.055, r = 0.304$), but showed a trend towards the extended UI. The results are illustrated in Figure 6.

Besides this general evaluation of the provided feedback, participants who tested the extended UI were asked to rate the usefulness of specific UI elements on a 5-point Likert scale. Since participants of the control group did not experience these UI elements, we asked them to rate the desirability of such additional cues instead. The resulting scores are listed in Figure 6.

Discussion

Both tested UIs achieved over-average usability scores, which is a point in favor of floor-based interaction techniques in SAR environments. However, in the concluding questionnaire six participants of the control group reported their confusion about who was able to control the scene elements, which is in line with the oral feedback after the study sessions. A reason why this difference between the UIs is not reflected in the usability scores might lay in the between-subjects design, since participants did not have any reference value. Regarding the feedback quality of the two tested UIs, we found an increased usefulness and accuracy of the extended UI. Particularly, both cues that were designed to support the interaction with the system, namely the footsteps and the progress bar, achieved the best ratings of all UI elements. Pre-tests revealed that footsteps are inevitable to be able to interact with the system on one’s own, especially for users who are not experienced in video games. We therefore decided to tell participants of the control
While both the buttons’ layout and the 2D floor maps were well received, the ROI segment attained the lowest score of all which percentage the participant’s view matched the intended view direction. Two additional floor-projected cues directed users to favorable viewpoints, the buttons themselves and scene-dependent 2D maps that emerged around the buttons. Both elements got neutral to positive reviews by users of the extended UI (see Figure 6).

User Guidance
To investigate to what extent the system allowed the users to follow the story, we set up a questionnaire that asked for four different aspects of storytelling, including the clarity of the storyline, the obviousness of where to look at, the feeling of disorientation and the fear of missing important story elements. Each item was measured on a 5-point Likert scale that ranged from Strongly disagree to Strongly agree. We ran four unpaired t-tests to analyze the questionnaire’s results, however, no significant effect was found.

In addition to the subjective rating of the storytelling we used an objective measure to evaluate the view direction of the participants during story-driven scenes. For every scene, regions of interest (ROIs) were defined that changed with regard to their size and position over the course of the story. While a floor-projected circular segment pointed to these regions in the test condition, they were invisible in the control condition. For each participant a ratio was calculated that describes to which percentage the participant’s view matched the intended view direction. The difference between the values of participants testing the basic UI ($M = 0.588, SD = 0.084$) and the extended UI ($M = 0.563, SD = 0.129$) was compared with an unpaired t-test, however, no significant effect was found.

Besides the ROI segment, which was designed to indicate a good view direction, two additional floor-projected cues directed users to favorable viewpoints, the buttons themselves and scene-dependent 2D maps that emerged around the buttons. Both elements got neutral to positive reviews by users of the extended UI (see Figure 6).

Discussion
While both the buttons’ layout and the 2D floor maps were well received, the ROI segment attained the lowest score of all feedback cues. Moreover, 4 of 20 participants of the test group did not notice the element at all. This was also confirmed by users of the basic UI. With an average score of $M = 2.90$ the support of storytelling through a floor cue was the least wished feature in a future UI. One reason for the low ratings may be the manageable size of the 4-sided CAVE, since all projection surfaces could be observed at once, either directly or peripheral. In addition, the story was rather simple with most story elements being presented in the center of the CAVE. Regions of interest inherently attracted the attention due to the movement within the scene. This assumption is also supported by the storytelling questionnaire, which resulted in scores around 4 out of 5 for both UIs. In addition, participants noted that they had to decide whether to focus on the floor-projected segment or the story, since it was difficult to bring both parts into view simultaneously. Therefore, future SAR UIs might want to do without floor-projected storytelling cues in favor of cues that are directly integrated into the scene.

User Collaboration
In order to analyze the master-follower concept, which was introduced to support user collaboration, we used both subjective and objective measures.

A questionnaire to measure the group accord was used as suggested in [35]. For each participant, we constructed an overall score from six questionnaire responses: the degree of enjoyment, the desire to meet the study partner again, the extent of perceived isolation, the degree of comfort with the partner, the degree of embarrassment induced by the partner, and the extent of perceived cooperation. Analysis of the results of the questionnaire with a two-way ANOVA did not reveal any significant effects. Group accord scores were similar for both the basic UI ($M = 80.278, SD = 14.247$) and the extended UI ($M = 77.778, SD = 12.998$). There was also no significant difference between unfamiliar partners ($M = 81.25, SD = 12.254$) and familiar partners ($M = 76.806, SD = 14.651$), although unfamiliar partners even reached a slightly higher score.

To gain further insight into the group behavior of users, we measured the distance between the users’ heads during the scenes and calculated a mean value for each pair.
We found a main effect of the type of UI on the mean utterances had to be discarded, because they were too.

The vertical bars show the standard deviation.

Figure 7: Pooled results of (a) the distance between the partners’ heads, (b) the categorized speech data, and (c) the balancing of the master and follower roles. The vertical bars show the standard deviation.

Table 1: Categories that were used to analyze the speech data.

<table>
<thead>
<tr>
<th>Speech Data</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interaction-related</td>
<td>Discussion or interpretation of interaction methods and UI elements.</td>
</tr>
<tr>
<td>Social or Emotional</td>
<td>Social or emotional utterances, such as laughing or an expression of excitement.</td>
</tr>
<tr>
<td>Technical</td>
<td>Hardware- or software-related discussions.</td>
</tr>
<tr>
<td>View-related</td>
<td>Discussions related to the perspective and visual perception of the participant himself or his partner.</td>
</tr>
<tr>
<td>Action-related</td>
<td>Planning of the own or the partner’s next actions.</td>
</tr>
<tr>
<td>Content-related</td>
<td>Discussion of elements, which are presented visually or auditory during the scenes.</td>
</tr>
</tbody>
</table>

of participants. Requirements for normally distributed data were fulfilled and the assumption of equal variances was not rejected by Levene’s test, so we ran a two-way ANOVA. We found a main effect of the type of UI on the mean head distance ($F(1,16) = 9.224, p = 0.008, \eta_p^2 = 0.366$), indicating a significant difference between users of the basic UI ($M = 1.575, SD = 0.421$) and the extended UI ($M = 1.177, SD = 0.237$). The familiarity also showed a main effect ($F(1,16) = 4.718, p = 0.045, \eta_p^2 = 0.228$), indicating a significant difference between partners, who knew each other ($M = 1.234, SD = 0.307$), and strangers ($M = 1.518, SD = 0.426$). No significant interaction effect was found between type of UI and familiarity ($F(1,16) = 3.801, p = 0.069, \eta_p^2 = 0.192$).

Besides logging the participants’ positions, we also captured their behavior on video. This allowed us to analyze the communication between partners. Based on [36], we considered utterances of the participants during the study. In comparison to the stated paper, we slightly extended the definition of an utterance to an individual word, sentence or even a small unit of a conversation between the participants, as long as their statements directly correlate. All utterances were assigned to the categories that are defined in Table 1. 10 of overall 763 utterances had to be discarded, because they were too low-voiced or slurred. The remaining utterances were analyzed using a two-way ANOVA with the type of UI and the familiarity of partners as independent variables. Although no significant effect of the overall communication could be found, the distribution of utterance types differed between the two UIs, as illustrated in Figure 7b.

To address the questions that are stated in section 3.3, we asked participants how they subjectively perceived the balancing of the master and follower roles as well as the master-driven leaving of a scene. Concerning the latter issue, 32 of the participants opted for the current solution, while only 7 participants would have preferred an automatic timer to leave the scene. One participant suggested to introduce a consensus mechanism. The opinions regarding the role assignment were divided. 23 of the participants decided for the current mechanism on a first-come-first-serve basis while 17 participants preferred an automatic assignment of the master role that is balanced between the users. However, only 8 of the 23 proponents of the first-come-first-serve technique were members of the control group.

In addition to this subjective evaluation, we also analyzed the actual balancing of roles between the two partners of a study session. For this purpose, we determined how often each partner held the master role. Afterwards, the ratio of both values was calculated by dividing the minimum by the maximum. Therefore, a balancing ratio of 1.00 corresponds to a perfectly balanced role assignment, while the balancing gets worse with lower ratios. The results, as shown in Figure 7c, were analyzed with a Mann-Whitney-U tests, since a normal distribution of the data cannot be assumed. We found a significant effect at the .05 significance level ($U = 14,000, p = 0.005, r = 0.620$).
As for user interaction and guidance, we also let the participants rate the UI elements, which were designed to support user collaboration. Participants of the test group rated the colored follower circle that pointed to the master with a mean score of $M = 3.80$, which is the third-best value of the six tested UI elements.

**Discussion**

During the study, we observed a highly collaborative behavior of the participants, both for familiar and unfamiliar partners as well as for both UI conditions. This impression matches the results of group accord scores and the measured amount of communication. However, the results also reveal differences in some measures, including the head distances, communication subjects and role balancing.

The mean head distance for participants using the extended UI was significantly smaller than in the basic UI, which indicates a positive effect of the floor-projected cues. Since users with the follower role were standing closer to the master, it can be assumed that they had more favorable viewpoints. Although no significant interaction effect between the UI type and the familiarity of partners on head distance was found, Figure 7a shows an interesting trend. While the head distance is almost the same for familiar and unfamiliar partners in the extended UI, unfamiliar partners kept a bigger distance from each other in the basic UI. This could indicate a positive effect of the extended UI elements on the convergence between users who meet each other for the first time.

Categorization of the speech data during the study revealed that users of the extended UI talked more often about the interaction with the system. This could be interpreted in two ways. On the one hand, users were more focused on the UI, since more elements that gave room for interpretation were present. Therefore, the participants might be more distracted from the actual content that was presented and therefore the learning outcome could be reduced. However, responses of the users of the extended UI also suggest a lower level of frustration while interacting with the system, which, on the other hand, could improve the learning experience. Future investigations should focus on the effects of different UIs on learning, before such systems can be used in an educational context.

Another significant difference was found regarding the user-driven assignment of the master and follower roles. While the distribution of roles was almost balanced between both partners in the extended UI, a high disparity could be observed in the basic UI condition. However, video inspection indicates that this is not caused by an unfair behavior of users, but by difficulties in understanding the interaction mechanism. Several groups who tested the basic UI had bad guesses on how to enter a scene, including the simultaneous standing of both partners on one or even two different buttons, the malfunction of one of the trackers, and the idea that only one user is qualified to be the master. Consequently, the partner who seemed to be tracked more reliable in the early stage of the study was chosen to be the master in a group consensus. This confusion could also be the primary reason why only 8 members of the control group voted for a user-driven balancing approach and the majority preferred an automatic, fair assignment of the roles instead.

**CONCLUSIONS**

In this paper, we presented a multi-user floor-based interface for SAR environments. We developed a number of floor-projected cues that aim to support users in several aspects when interacting with such systems. To evaluate the effectiveness of the cues, we performed a user study with 20 pairs of participants. The results indicate that the interface is self-explanatory and easy to use, and therefore could be used in public environments such as museums without the need for support of an additional instructor. It also fostered communication between partners, both between friends and strangers, and encouraged users to move closer together. By this means, better viewpoints could be ensured for multiple users. From a narrative perspective we could not observe significant improvements, however, high scores were achieved even in an application without additional feedback cues. When faced with a more complex story, users might need more support to keep track of where to look at within the scene. Furthermore, future studies should focus on the learning success achieved by the system. After the present study it is still an open question whether the floor interface directs users from the presented content or if it sparks the users’ interest in a new topic. Further investigations could pave the way for floor-projected UIs to be used in real scenarios.

**REFERENCES**


