

Audio-Visual Attractors for Capturing Attention to the Screens When Walking in CAVE Systems

Francesco Grani, Ferran Argelaguet Sanz, Valérie Gouranton, Marwan Badawi, Ronan Gagne, Stefania Serafin, Anatole Lécuyer

► To cite this version:

Francesco Grani, Ferran Argelaguet Sanz, Valérie Gouranton, Marwan Badawi, Ronan Gagne, et al.. Audio-Visual Attractors for Capturing Attention to the Screens When Walking in CAVE Systems. Sonic interaction with a virtual orchestra of factory machinery, IEEE VR, Mar 2014, Minneapolis, United States. pp.1-4. hal-01018818

HAL Id: hal-01018818

<https://hal.inria.fr/hal-01018818>

Submitted on 7 Jul 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Audio-Visual Attractors for Capturing Attention to the Screens When Walking in CAVE Systems

F. Grani, S. Serafin*
Aalborg University Copenhagen

F. Argelaguet, V. Gouranton, M. Badawi, R. Gaugne, A. Lécuyer†
Inria-IRISA, Rennes

ABSTRACT

In four-sided CAVE-like VR systems, the absence of the rear wall has been shown to decrease the level of immersion and can introduce breaks in presence. In this paper it is investigated to which extent user’s attention can be driven by visual and auditory stimuli in a four-sided CAVE-like system.

An experiment was conducted in order to analyze how user attention is diverted while physically walking in a virtual environment, when audio and/or visual attractors are present. The four-sided CAVE used in the experiment allowed to walk up to 9m in straight line. An additional key feature in the experiment is the fact that auditory feedback was delivered through binaural audio rendering techniques via non-personalized head related transfer functions (HRTFs). The audio rendering was dependent on the user’s head position and orientation, enabling localized sound rendering.

The experiment analyzed how different “attractors” (audio and/or visual, static or dynamic) modify the user’s attention. The results of the conducted experiment show that audio-visual attractors are the most efficient attractors in order to keep the user’s attention toward the inside of the CAVE. The knowledge gathered in the experiment can provide guidelines to the design of virtual attractors in order to keep the attention of the user and avoid the “missing wall”.

Index Terms: Audio user interfaces, head-related transfer function, 3-D audio, virtual auditory space, navigation, walking, virtual reality, multimodality, virtual environments.

1 INTRODUCTION

In a four-sided CAVE environment, the absence of the rear wall is prone to decrease the user immersion and introduce breaks in presence during a virtual experience [11]. Although real walking is often the most appropriate navigation interface in virtual environments, the walking area is restricted by the physical laboratory space. Several solutions have been proposed to cope with the limitations of a laboratory’s physical space, such as redirected walking [2], redirection based on change blindness [12], redirected walking in place [9, 8] and natural metaphors [3].

In this paper, we investigate a different approach, the usage of virtual attractors in order to capture users’ attention away from the missing wall. While studies in auditory perception and cognition have demonstrated the importance of sound to capture users’ attention in real environments [5, 10] there is still a lack of research in Virtual Environments. An exception to this lack of studies in audio-visual attractors in VE is the work presented in [6, 7], where reorientation techniques based on the concept of visual and auditory distractors were proposed. Results showed that an audio distractor did not produce as high a feeling of presence as a natural

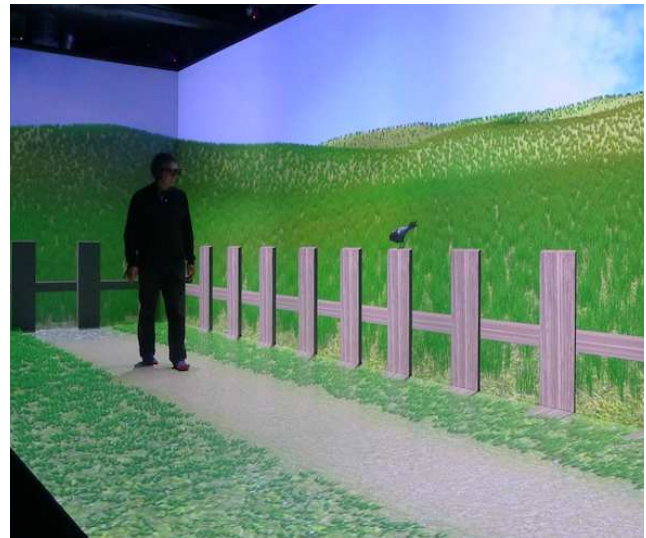


Figure 1: A user walking in the wide four-sided CAVE towards one of the short sides. The attractor (virtual crow) is placed on the front screen of the CAVE in order to divert the users’ attention.

audio-visual distractor. However, it produced a higher feeling of presence than an unnatural distractor without audio.

In this paper, we present an experiment whose goal was to investigate whether the use of auditory and visual attractors captures users’ attention, in such a way that they do not notice the missing wall in a four-sided CAVE system. We combined several audio and visual attractors and analyzed the user behavior in several conditions. The results showed that the user behavior while physically walking can be modified by the addition of audio-visual attractors inside the virtual environment.

2 EXPERIMENT

The goal of the experiment was to explore how users’ attention is driven by visual and auditory stimuli in a four-sided CAVE. In the experiment, we employed visual and auditory attractors and measured to which extent users’ attention was modified by the presence of such attractors while walking in a virtual environment. The experiment was conducted in a wide four-sided Virtual Environment which enabled users to physically walk along a path of 9 meters (see Figure 1).

Two groups of attractors were considered: “bad” and “good” attractors. “Good” attractors are virtual objects placed in the projected space opposite to the missing wall: these can be either visual, audio-visual, or auditory stimuli. In the contrary, “bad” attractors those placed outside the CAVE by the missing wall side; these kind of attractors can only be auditory. In order to have the same configuration on both the directions of walking, “good” and “bad” attractors were placed in a symmetrical configuration with respect to the longitudinal axis of the CAVE (coincident with the path to be walked by users) to explore how users’ behavior was modified. See

*e-mail: {fg, sts}@create.aau.dk

†e-mail: {valerie.gouranton, ronan.gaugne}@irisa.fr
{fernando.argelaguet_sanz, marwan.badawi, anatole.lecuyer}@inria.fr

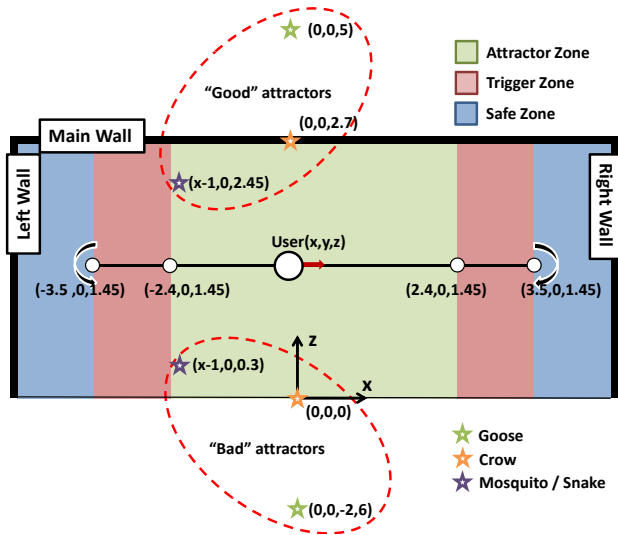


Figure 2: Schematic top view of the experiment layout indicating the CAVE walls, the testing path to be walked by the subjects, and the positions of the different Attractors.

Figure 2.

2.1 Apparatus and Participants

The experiment was run in a four-sided CAVE with retro-projected glass screens. The dimensions of the CAVE are: Main screen: 9.6m x 3.1m (12.5 MPixels); lateral screens: 3m x 3.1m (2.3 MPixels each); horizontal acrylic ground: 9.6m x 2.9m (3.6 MPixels). Users' head position and orientation was tracked by 16 ART Tracking infrared cameras, using reflective markers placed on the 3D glasses. Finally, audio rendering was delivered using Sennheiser HD600 full-sized circumaural open-air headphones and a wearable radio receiver (to avoid cabling issues in the walking task).

Thirteen subjects participated in the experiment (4 Female, 9 Male). Aged between 21 and 59 ($\bar{x} = 31.3; \sigma = 9.7$). Each participant was able to finish the experiment in less than 45 minutes.

2.2 Visual rendering

The environment chosen for the experiment was a natural VE without salient virtual objects (e.g. tree, rocks). This was made in order to limit salient features which might have driven the attention of the user. As users were instructed to walk along the CAVE, the path the users had to follow was encoded in the ground texture of the environment (gravel trail). For safety reasons, a virtual fence was introduced in order to avoid the user to run through the walls of the CAVE and also to provide additional cues regarding the path-following task (clearly defining the limits). See Figure 1.

Unity3D and Blender were used to design both the visual environment and the visual attractors while MiddleVR was used to drive the CAVE rendering, synchronization and the tracking system.

2.3 Auditory rendering

Binaural audio rendering was performed in real-time on a MacBook Pro running MaxMSP 6 and *High Order Ambisonic Library 1.2* released from CICM - Maison des Sciences de l'Homme Paris Nord [4]. This library first encodes the soundfield in High Order Ambisonic to render it then to different output possibilities (different speakers layouts, headphones). The *hoa.binaural* object contained in the library virtualizes ambisonic soundfield using HRTF to perform FIR filtering of the incoming signal. HRTFs are extracted from the CIPIC HRTF database [1].

The relative coordinates of all the sounding objects (with respect to the users' head position and orientation) are sent from Unity3D to MaxMSP as raw data packages via UDP to achieve a minimum latency in the binaural audio rendering. The sound materials used to create the natural audio landscape as well as the animal sounds were Creative Commons sources gathered from the Freesound.org archive.

2.4 Procedure

The task to be performed by the user was simple, walk along a straight path, parallel to the main wall of the CAVE. The path was clearly visible for the user as it was represented on the floor as a gravel trail. After explaining the task to the participant, the examiner asked him to step inside the CAVE and place himself in one of the sides. Users were then instructed to walk straight without stopping.

Participants started each trial from the Safe Zone (see Figure 2) and walked straight to the Safe zone on the other side of the CAVE. When entering the Attractor Zone, the current virtual attractor was enabled (according to the current condition). Once the user covered the Attractor Zone and reached the second Trigger Zone, the attractor was deactivated. Finally, when reaching the end of the path in the Safe Zone participants had to perform a u-turn and repeat the task until the end of the experiment. The end of experiment was indicated by a virtual stop sign which appears in the center of the CAVE.

2.5 Design and Hypotheses

The experiment covers different configurations of virtual attractors. In this study, we are considering the *Stimuli* (Baseline, Auditory, Visual and Auditory), the *Position* of the stimuli (inside/outside the CAVE) and the *Relationship* between the stimuli and the user (fixed position, or following the user). For practical reasons, we constrained the number of levels for the *Relationship* factor. We only considered two fixed attractors (Goose and Crow) and two moving attractors (Mosquito and Snake). Figure 2 depicts the position of the attractors with respect the user and the CAVE, note that the position is symmetrical with respect to the path. Furthermore, as environmental factors (physical environment) or user factors (eye dominance, handedness) can influence the user behavior when performing the task, we considered the *Direction* of the walk as an additional factor with two levels: Left to Right (L2R) and Right to Left (R2L).

Regarding the factors *Type* and *Position* due to obvious conflicts (e.g. when the attractor is placed outside the CAVE it can not be visible) we merged them in a new factor *Condition* with the levels: Baseline (no attractors), Audio Inside (AI), Audio Outside (AO), Audio + Visual (AV). This lead to factorial design of 4x4x2 (Condition x Attractor x Direction) resulting in 32 combinations. To decrease the data variability we considered 3 repetitions for each combination, leading with a total of 96 trials presented in random order. In addition we included 4 training trials (with no attractors) were added at the beginning of the experiment to ensure that the participants understood the procedure.

The dependent variables of the experiment were (1) the mean orientation of the users' head while performing each trial and (2) the mean speed of the user per trial. For the analysis, only orientation data recordings inside the Attractor Zone area are considered (see Figure 2).

In the experiment we wanted to explore how the different stimuli influenced the user behavior. Our hypotheses considering the design of the experiment were:

- H1 The mean users' orientation for the "Baseline" condition will be zero.
- H2 Users will be more attracted towards the attractor when the attractor is visible and audible.

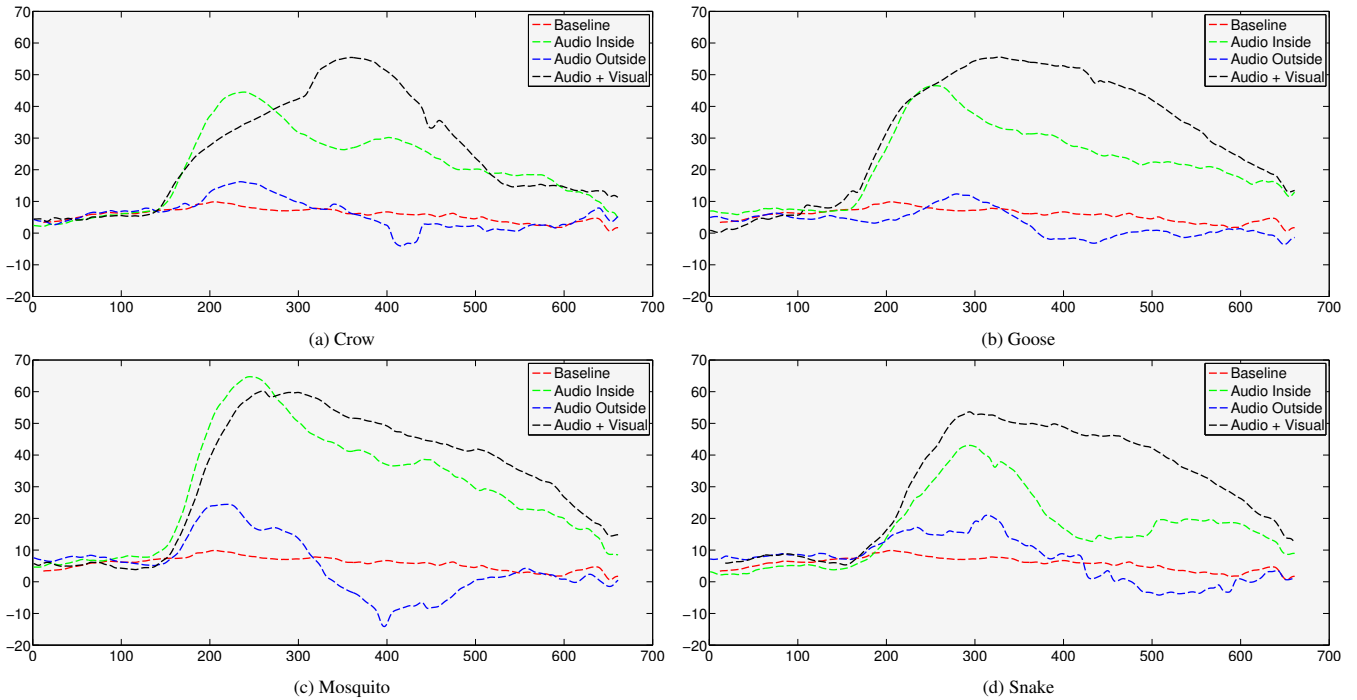


Figure 3: Average mean orientation of users’ head along the path for each attractor. X-axis is centimeters, Y-axis degrees. Positive angles means that users is looking towards the Main Wall. We observe the immediate reaction of the user when entering the Attractor Zone (150cm). It is interesting to notice users’ behavior for the mosquito in the Audio Outside condition. First, users tried to locate the mosquito in the Main Wall and then shifted their attention towards the “missing wall”. This effect is also visible for the other attractors although it is weaker.

- H3 The mean head orientation will be higher when the attractors are inside the CAVE.
- H4 The mean head orientation will not be dependent on the *Direction*.
- H5 The mean head orientation will be higher for moving attractors (Mosquito, Snake).

At the end of the experiment participants had to fill a questionnaire in order to gather subjective information about the level of attractiveness of each attractor considering each type of feedback (7-Likert scale questionnaire).

2.6 Results

The analysis of the mean speed did not showed any significant differences among the different factors. In average the mean user speed was of $0.95m/s^2$ with a $\sigma = 0.16m/s^2$.

2.6.1 Mean Orientation

In order to analyze the head orientation data, we computed the mean orientation for each trial. As a convention, if the orientation is positive, the user is looking towards the inside of the CAVE. A value of 0 is that the head orientation is parallel to the main screen.

We first checked the validity of H1. For this purpose we performed a t-test checking whether the mean head orientation for the trials in which no attractor was present was equal to zero. The t-test rejected the NULL hypothesis ($p = 0.001$), thus we have to reject H1. The mean head orientation is different than zero. Participants were attracted to the main wall of the CAVE although any attractor was active.

We conducted an analysis of variance (ANOVA) to analyze the effect of *Condition* and *Direction* in the mean head orientation. For the post-hoc tests we used the Bonferroni method with an ($\alpha = 95\%$), only significant pairwise comparisons are discussed (all $p < 0.05$). The Anderson-Darling normality test showed that the data was normally distributed ($p < 0.001$). The two-way ANOVA

of the mean orientation of the user versus the *Condition* and *Direction* showed two main effects: (1) there is a significant difference among *Conditions* ($F(3,36) = 154.75; p < 0.001$) and (2) there is a significant difference among *Directions* ($F(1,12) = 97.30; p < 0.001$). Post-hoc tests showed that among the different *Condition* the AV condition was the one driving more the users’ attention (thus accepting H2) followed by the AI condition. The AO and the Baseline conditions were the conditions driving less the users’ attention (there is no significant difference among them), thus accepting H3. Post-hoc tests among the two levels of *Direction* showed a significant mean difference among the two levels, R2L ($\bar{x} = 20.887; \sigma = 21.931$) and L2R ($\bar{x} = 11.070; \sigma = 21.730$). Users performing the task walking from the right side of the CAVE towards the left side, showed a higher mean head orientation, thus rejecting H4.

In order to analyze the effect of the attractors (Bird, Goose, Mosquito and Snake), we performed a second analysis removing all the trials for the Baseline condition. The two-way ANOVA of the mean orientation versus *Condition* and *Attractor* showed two main effects: (1) a significant effect on *Condition* ($F(2,24) = 128.66; p < 0.001$) which is consistent with the previous analysis (same pairwise differences) and (2) a the significant effect on *Attractor* ($F(3,36) = 3.39; p < 0.05$). Post-hoc pairwise tests for the four levels of *Attractor* showed only a significant pairwise difference between the mosquito ($\bar{x} = 22.88; \sigma = 28.37$) and the crow ($\bar{x} = 17.60; \sigma = 19.82$) attractors.

2.7 Questionnaires

A post-experimental questionnaire was given to the subjects, asking the degree to which they noticed the role of animal sounds to capture attention in all three conditions (see Figure 4).

A two-way ANOVA of attractiveness vs *Condition* and *Attractor*, showed two main effects on *Condition* ($F(2,24) = 14.51; p < 0.001$) and *Attractor* ($F(3,36) = 6.07; p < 0.001$). In addition, it

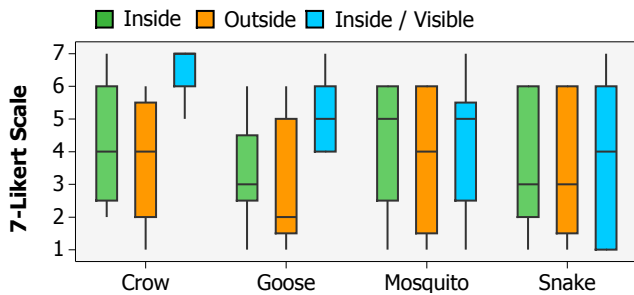


Figure 4: Level of attractiveness (subjective questionnaires) for each attractor. Plot shows a clear increase of attractiveness for the crow and the goose when visible and audible at the same time.

also showed an interaction effect between both factors ($F(6, 72) = 2.89; p < 0.05$). Post-hoc analysis showed that attractors combining sound and visual feedback had a level of attractiveness significantly higher ($p < 0.05$). Also, regarding the different attractors, the Crow also presented a significantly higher level of attractiveness ($p < 0.05$), thus rejecting H5. The interaction effect was related to the fact that the attention captured by the attractors inside the field of view of the user (goose and crow) is greatly increased when visible and audible at the same time.

2.8 Discussion

The results showed that the addition of attractors in the virtual environment is able to modify the user behavior when physically walking. The results of the mean head orientation while walking (see Figure 3) showed that the effect is stronger when the virtual attractor was easily localizable by the user (Audio + Visual attractor). Also, we have to notice that even in the absence of attractors users are slightly attracted towards the inside of the CAVE. If we observe the behavior of the user along the path, we can clearly observe how the user orientation is modified along the trajectory for all attractors and also the fact that when the attractor is visible and audible, they are able to keep the attention of the user during a longer period of time. In addition, we can observe the clear difference between the Audio Inside and Audio Outside conditions. “Good” attractors are able to drive the users’ attention towards the inside of the cave while the “bad” attractors, in which the audio source was placed outside the cave, are not able to modify the user’s behavior. This result is also supported by the subjective questionnaires.

Another interesting (and unexpected) result was the fact that a difference in behavior appeared according the direction of the walk. A possible explanation for this behavior could be also the presence, by the missing wall, of other kinds of “Attractors” such as physical stimuli present in the lab space (hardware systems, a door, the operator).

Finally, the higher level of attention towards the crow in the questionnaires contradicts the analysis of the mean orientation and our hypothesis. However, several facts can explain the lower mean orientation (1) the Crow was the attractor closest to the user (2) when the Crow appears, it is already inside the field of view of the participants and (3) it was the easiest attractor to locate. In Figure 3 we can observe how the changes of head orientation are smaller while approaching the Crow (from 150cm to 300cm) compared to other attractors which resulted in stronger orientation changes.

3 CONCLUSIONS

In this work, we have explored how user attention is modified when virtual attractors are included in the virtual environment. Among the different auditory and visual stimuli analyzed, attractors combining visual and auditory cues provided the best results. From the

results obtained, we can affirm that audio-visual attractors are able to drive the user attention towards the inside of a CAVE-like environment. Another important result is that attractors (audio sources) placed outside the CAVE are not able to modify the user behavior (as there are no significant differences compared with the Baseline condition) and that for our particular setup the direction of the walk influences the user’s behavior. While the first and second results matched our hypotheses, the third is unexpected. Possible explanations for this unexpected result are users’ eye-dominance or the physical stimuli present in the lab. Both factors could have biased the experiment. Additional experiments need to be performed to confirm our hypotheses.

The results we are presenting here could be promising for a wide range of applications as the audio-visual attractors could be employed to attract users’ attention and could be adapted for its use to a variety of applications and tasks. For example, results could be applied to avoid users getting too close to the walls in a CAVE-like environment and in general for redirection and navigation purposes in games, architecture exploring, virtual prototyping and many other applications.

As future directions, we want to extend the analysis to explore a wider range of attractors (e.g. virtual avatars, task-specific attractors), analyze how user attention can be diverted when the user is performing a task involving a higher mental workload or how the approach can be extended to other VR setups. Finally, the ultimate goal could be the design an algorithm capable of generating automatically attractors when detecting that the user is about to look towards an unwanted direction.

REFERENCES

- [1] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano. The cipic hrtf database. In *IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, pages 99–102, 2001.
- [2] G. Bruder, F. Steinicke, and K. H. Hinrichs. Arch-explore: A natural user interface for immersive architectural walkthroughs. In *IEEE Symposium on 3D User Interfaces*, pages 75–82, 2009.
- [3] G. Cirio, M. Marchal, T. Regia-Corte, and A. Lécuyer. The magic barrier tape: a novel metaphor for infinite navigation in virtual worlds with a restricted walking workspace. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, pages 155–162, 2009.
- [4] J. Colafrancesco, P. Guillot, E. Paris, A. Sèdes, and A. Bonardi. La bibliotheque hoa, bilan et perspectives.
- [5] J. B. Fritz, M. Elhilali, S. V. David, and S. A. Shamma. Auditory attention-focusing the searchlight on sound. *Current opinion in neurobiology*, 17(4):437–455, 2007.
- [6] T. C. Peck, H. Fuchs, and M. C. Whitton. Evaluation of reorientation techniques and distractors for walking in large virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 15(3):383–394, 2009.
- [7] T. C. Peck, H. Fuchs, and M. C. Whitton. Improved redirection with distractors: A large-scale-real-walking locomotion interface and its effect on navigation in virtual environments. In *IEEE Virtual Reality*, pages 35–38, 2010.
- [8] S. Razzaque et al. *Redirected walking*. University of North Carolina at Chapel Hill, 2005.
- [9] S. Razzaque, D. Swapp, M. Slater, M. C. Whitton, and A. Steed. Redirected walking in place. In *Proceedings of the workshop on Virtual environments 2002*, pages 123–130. Eurographics Association, 2002.
- [10] B. G. Shinn-Cunningham. Object-based auditory and visual attention. *Trends in cognitive sciences*, 12(5):182–186, 2008.
- [11] M. Slater and A. Steed. A virtual presence counter. *Presence: Teleoperators and virtual environments*, 9(5):413–434, 2000.
- [12] E. A. Suma, S. Clark, D. Krum, S. Finkelstein, M. Bolas, and Z. Warte. Leveraging change blindness for redirection in virtual environments. In *IEEE Virtual Reality*, pages 159–166, 2011.