Journal of Applied Mathematics

volume III (2010), number I

THE INFLUENCE OF DYNAMIC SCENE CHANGES IN GEOMETRY OF VIRTUAL POSITIONING TASKS

PETRASOVA Alena, (UK), CZANNER Silvester, (UK), CZANNER Gabriela, (UK), FARRER Valerie Julie, (UK), CHALMERS Alan, (UK), WOLKE Dieter, (UK)

Abstract. In this paper we investigate how different conditions aiding perception in an online Virtual Reality application affect depth perception. The VR was developed for parents experiencing feeding difficulties with their infants. The role of effective use of stereoscopic viewing, head tracking and animation is examined. In a complete block designed experiment subjects were asked to determine the positions of the spoon in the 3D environment. Task was done under different conditions: the animation of the baby model was either on or off and four different viewing modes were used: monoscopic or stereoscopic viewing with or without head tracking. The use of head tracking improved the position perception when monoscopic vision without any animation applied. However, adding head tracking to stereo vision did not improve the position perception. The use of animation positively affected the judgment of depth perception under the monoscopic vision and makes it comparable with the performance under the head tracking mode. The outcome of this study provides insight into how to create more interactive scene and enhance depth-related visual tasks by exploiting dynamic scene changes.

Key words. 3D navigation, visual perception, virtual reality, virtual characters, user studies, feeding disorders.

Mathematics Subject Classification: Primary 68N30; Secondary 62-07, 62K10.

1 Introduction

The possible impact of Virtual Reality (VR) on health care is significant because it is not only a communication interface but also an experience and its key characteristic is the sense of presence [1]. Virtual reality can be described as a technology or tool for influencing cognitive operations. The users learn to consider different interpretations of a situation and they develop they own list of problem situations, which they discuss with the therapist and he makes the decision on how to proceed next. Cognitive behaviour therapy (CBT) is an approach based on modifying distorted beliefs, attitudes and cognitive processes that maintain disordered behaviour [2]. We aim our attention to observe the interaction between mother and child, because the quality of the parent-infant interaction may affect the quality of action during feeding [3].

The main goal of this work is to investigate what would be the best affordable and effective VR environment for feeding simulation purposes. A unique virtual environment application is being developed for individualised interactive therapy incorporating components of CBT, which is built on various information, such as the right environmental conditions for meal times, eating distractions, the behaviour and social skills of the child, suggested by physicians and psychologists, which may enable parent's successful experience with our application. According to Sherman and Craig [4], VR as a medium of communication requires physical immersion and interaction for essential experience. The potential advantage for participants is their ability to impact the narrative flow of the experience. Our motivation for developing the application is to significantly improve a new parent's understanding and treatment of infants with feeding disorders. To achieve this we are investigating a novel online Virtual Reality solution. For effective interaction with the virtual baby, it is necessary to concentrate, besides the emotional reactions and expressions of the character, on users' perception of the space and relative positions of individual objects.

2 Related Work

Based on the previous experiment by Petrasova et al. [5], we found users having some problems to perceive depth and spatial information in our virtual environment. Several times users were confused with the model of a spoon near the model of a baby and they were not able to describe the spoon's actual position in relation to the baby model. As the most common users' problems lie in the poor estimation of distance in a virtual 3D scene, there is a number of observations of differences in the perception of distance in real and virtual scene [6], [7]. There appears to be a significant effect on the accuracy of judging the perception of distance, with and without the perception of self in the environment [8]. Perception of absolute egocentric distance in VR environment was investigated also by Gooch and Willemsen [9]. They suggested perception of distance in immersive environment is effected by the display device as the distance judgment in photographic environments and environments modelled by traditional computer graphics is very close [9].

It is still our goal to understand how to successfully interact in the 3D space of virtual environment to help perform tasks. This is particularly important considering applications for training and simulation of everyday problems and cognitive tasks. Developing virtual environments that support learning requires an understanding of the relationships between the cognitive capacities of the learner, the environment and activities being modelled, and the technical affordances of VR that support information encoding, development of knowledge structures, and performance [10]. There are several studies, which deal with the perception of spatial information in virtual environment and perceptual accuracy and how to improve them. Bigoin et al. [11] investigated the correlations between the perception of depth and the feeling of discomfort by immersing participants in two different stereoscopic virtual landscapes. Correlation between depth perception and geometric characteristics of the virtual environment, such as sharp edges, was found. Higher depth sensation did not correlate with a higher level of discomfort, but in combination with the navigation in the virtual environment it may lead to bigger sense of immersion. There are not many observations related to the depth perception in the interaction with objects in the environment in terms of contact of objects. The effectiveness of binocular stereo, shadow and diffuse inter-reflection cues on the precision for imminent contact with object was evaluated by Hu et al [12]. It was found that subjects were more precise with binocular disparity than without it [12]. Binocular stereopsis seems to have also impact on our ability to judge the visual realism of computer generated images [13]. This is expected due to the fact that stereo viewing is closer to our natural viewing experience.

More experiments are performed with the intent of determining how the perception of depth is caused by generating of shadows, lighting, textures or reflections of individual objects [12, 14 and 15]. There are many factors that may improve or influence the sense of spatial orientation, assessing of the size, mutual position of objects and precision in manipulating with virtual objects in the environment [16]. Each of them offers its own estimation of depth.

Motion pictures also allow us to better imagine objects shown on the screen as 3D models so they do not look flat. Based on the ability of brain to process visual stimuli the speed, direction and the placement of object in 3D space is analysed. Depth perception is a feeling extracted from flat stimuli [17] as the images cast onto our retinas varies in only two spatial dimensions [16]. Motion parallax can thus, with wise using, improve the impression of relative distance in the scene [18]. According to Mather [16], the optic flow and motion parallax are sufficient to support high-precision depth perception even when no other depth cues are available. However Jones et al. [19] found no motion parallax effects on depth judgment in VR environment, although their participants were wearing a head mounted display (HMD).

3 Experimental Conditions and Hypotheses

The study was conducted with two experimental conditions. Tasks were done with and without animation (dynamic and static tasks) and four different viewing modes were used - only laptop display (monoscopic vision) with full HD resolution (D1), stereo vision (D2), head tracking (view dependent rendering) with monoscopic vision (D3) and head tracking with stereo vision (D4). Our main interest was to observe the influence of individual experimental conditions on the better task performance and spatial perception of virtual environment. To maximize the effect of viewing mode and animation of the object on depth perception, cast shadows, which improve the perceived position of objects in 3D space, were not used. Only attached shadows were used, which provide powerful cues to 3D object structure [16].

To enhance the perception of 3D objects in the environment red-cyan anaglyph stereo pictures were used. Although it is known that this filter technique distorts the perception of colour [20], it was sufficient to represent the 3D scene. Several participants were observed at trying to catch the virtual spoon "leaving the screen", while they were using head tracking with the stereo vision mode. This technique is a simple and inexpensive way for an observer to see the stereogram. Stereo has certainly significant benefit and has been extensively supported by research. However stereo in combination with head tracking has not been studied thoroughly. Most of studies used HMDs, when the participant could at most rotate, but not tilt closer to the objects in an environment [9]. In contrast to other experiments aimed for investigating the influence of depth cues combination on the user's perception, for example [12], we allowed users to move towards and away from a model of the baby and also tilt from side to side. We propose the following hypothesis: *Head tracking will improve the depth and distance perception in virtual environment.*

A common problem with HMD, even it is updated quickly and accurately in response to head motions, is the remaining subtle delay causing desynchronization of motion parallax from the head motion and therefore weakening it as a depth cue [21]. The effect of object movement is therefore taken into account with regard to static user or the user moving relative to the scene. As in this experiment the combination of glasses that do not need to update was used, the following four factors when using the different display techniques were considered: the user and the baby model in the virtual environment were static, the baby model or participant is moved and the last one allowing the user and the baby model movement at the same time. We propose the following hypothesis: *The animation of the object of interest in Virtual environment will affect our judgment of depth perception.*

36 subjects participated in the experiment. There were 14 females and 22 males in the mean age 30.5 (19 to 64 years old). All of them had normal or corrected to normal vision with the use of glasses or contact lenses. After the experiment participants were asked to fill out a questionnaire. The participants rated their familiarity with computer games and graphics 2.5 in average on the 5 levels Likert scale with 1 representing familiar and 5 representing unfamiliar. They were also asked to rate their feeling of the sense of depth and distance in the VR application on the Likert scale of 5 levels with 1 representing strong and 5 representing poor. The mean rate of depth sense was 2.6.

4 Experimental Tasks and Their Performance Measures

Participants were presented with three different positions of spoon in relation to the model of a baby and the model of a bowl. The first position was under the baby's head and left cheek (from user's point of view) and not touching the baby's body, the second was in front of the baby in the bowl and nearer to the baby's left arm (Figure 1), the third was in front of the baby's head, next to its right cheek and not touching its body. Subjects were not aware that these three positions of the spoon are repeated randomly all the time. The task was to choose one of four options for each case describing where the spoon is situated in relation to the baby. The answers were rated according to their correctness from 1 meant correct to 4 meant wrong answers. There was always only one correct answer and one was a very close. The other two options meant greater error in the spatial perception and position of objects.

First two position examples of the spoon in the positioning task

Figure 1 on the right side depicts this task with monoscopic vision mode using one of the positions of the spoon. In this case the correct answer (score 1) is that the spoon is not touching the baby's body, it is under the baby's head and left cheek (from user's point of view). Because the spoon did not touch a baby model but was very close, the close to good response was the similar answer with the spoon is touching baby option and such an answer would yield score 2. Score 3 was for the answer, where the spoon is touching the baby's arm and is placed under its head, next to the left cheek. The example of a very wrong answer (score 4) is that the spoon is not touching the child , it is in front of the child, next to the left cheek.

Positioning task completion time was in the most cases up to 5 seconds. In general, it took longer time to choose the answer with the monoscopic vision viewing mode in compare to other display techniques. The task was constructed in our application, which is developed in the JMonkey game engine [22]. Stereoscopic viewing was enabled by using red-cyan paper glasses. Head tracking was provided by wearing safety glasses with two infrared LEDs (one at each side of glasses). Inspired by Lee's [23] program for head tracking for desktop VR displays using the Wii remote, we implemented our own library to the game engine. The relative position of two LEDs was observed by built-in camera of the Wii remote controller placed in front of the user.

5 Experiment Procedures

We observed the perception using 4 (viewing mode) × 2 (animation) × 3 (3 positioning tasks) conditions, resulting in each subject experiencing 24 trials during the complete block designed experiment. We analyzed the positioning-task data via Linear Mixed-Effect Models (MIXED) (also known as multilevel regression or mixed effects regression models) in SPSS 17.0. This is estimated via restricted maximum likelihood. This approach takes into account the multilevel structure of the data: for each participant there were 24 observations. We assumed random effect of the subjects and fixed effect for the viewing mode and task order of the task. We also controlled for the effect of the order of the viewing mode and for the effect of display position. If we found a significant effect of a factor we then performed mean comparison correction to ensure the 95% simultaneous confidence. MIXED models expand on the analysis of variance models have previously been used within perception studies by Mastoropoulou et al. [24], and are an extension of the repeated measures analysis of variance model used in positioning and resizing-task trials of Hubona et al. [15]. To judge the effect of animation on perception in positioning tasks we performed MIXED model analysis on the differences of the scores from the static and dynamic positioning tasks.

No strong evidence was found of the effect of the display order on the perception in the positioning task without animation (p-value = 0.185) neither with animation (p-value = 0.516). The order in which the positioning tasks are presented in the scene without the animation is important (p-value=0.007) only in display mode, however this order has no effect on the perception in none of viewing modes in the scene with animation (p-value=0.360).

6 Results

First, we investigated the differences in positioning tasks across the viewing modes. The subjects obtained the best (lowest) average score at viewing modes D2 and D4, worse at D3 and worst at D1 (Figure 2). Further, the number of participants who correctly determined the spoon position (i.e. received the score 1) also varied across viewing modes without the animation of the baby model. In the positioning tasks 1 and 3 (the spoon is close to the baby's face) 24 to 29 subjects correctly determined the spoon position at the viewing mode D2, then D4, whereas only 15 to 17 subjects correctly reported the position at the viewing modes D3 and D1 (Figure 2). In the positioning task 2 (the spoon in the bowl) the subjects tended to report the position of the spoon better than at other modes especially when head tracking is used.

The observed perceptual differences across the viewing modes (Figure 2) of the 36 subjects were found to be significant (p-value<0.001 in F-test of MIXED model). The order of displays from the best to the worst score for the static positioning task is: D2 = D4 > D3 > D1, where we use "=" to denote statistically equivalent conditions, and ">" to denote significant preference of the condition on the left side to the condition on the right side. Here, the perception at D2 and D4 is statistically same (p-value=0.999). The perception at D2 is significantly better than at D3 and D1 (p-values=0.006 and 0.001). The subjects achieved almost significantly better results using the combination of stereo vision with the head tracking than only with head tracking mode (p-value=0.013) and significantly better than at monoscopic vision (p-value<0.001). The use of head tracking is also significantly better than monoscopic viewing mode (p-values=0.016).

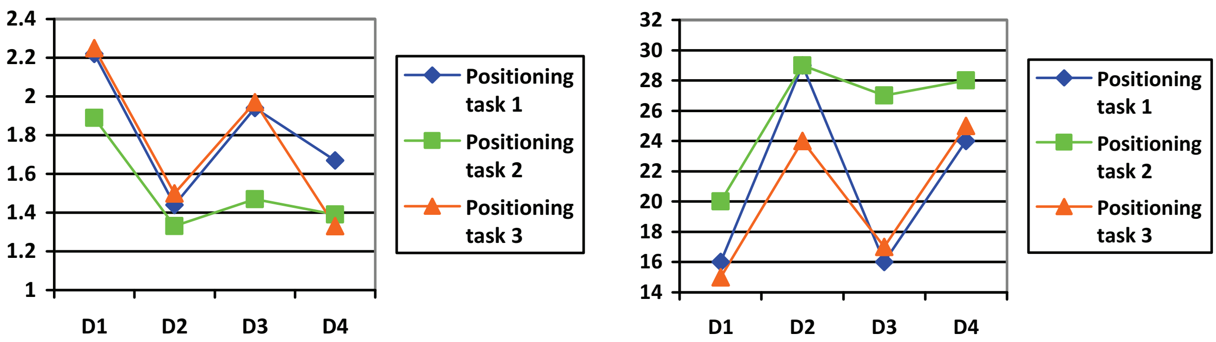

Figure 2. Average score in the three positioning tasks (on the left) and number of subjects, who correctly reported the spoon position (on the right) across the four viewing modes without animation.

Next we studied the positioning task in three types of the task. In our 36 subjects the average scores of tasks 1 and 3 (in both tasks the spoon is close to the face) were similar and worse than with task 2 (spoon in the bowl). Similarly, the number of subjects who correctly specified the position is the highest with task 2. Figure 2 suggests a different effect of the task across viewing modes, however this interaction was found to be insignificant (p-value=0.643). The differences in position perception across the types of tasks were found to be significant (p-value=0.025). There is evidential improvement of position perception in task 2 using head tracking when compared to the other two positioning tasks (p-values 0.011 and 0.037 in Sidak multiple comparisons test). We suggest that it may be caused by slightly different position of the spoon. In this task, the user needed to consider the position of the spoon in relation to a bowl in front of the model of baby (Figure 1, picture on the right) and the spoon was placed closer to user, so they had better opportunity to observe its position with head tracking than in the case of two other positions. The position perception at tasks 1 and 3 was found to be not significantly different (p-value 0.633 in Sidak multiple comparisons).

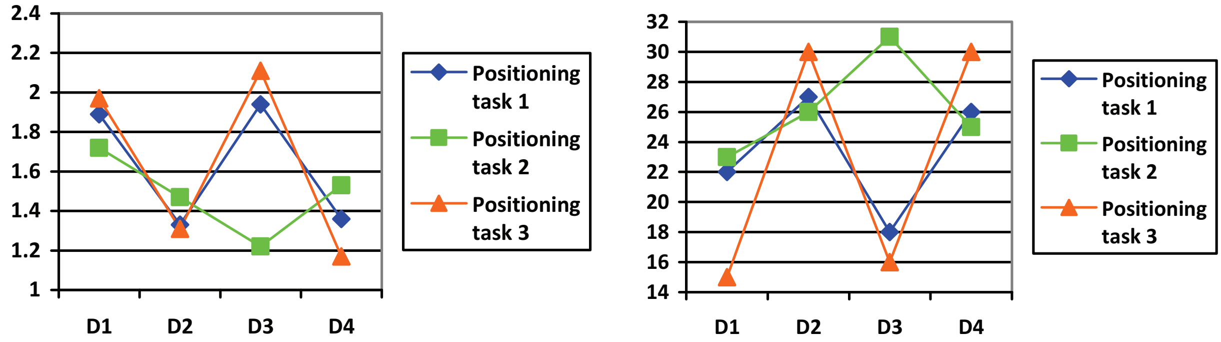

Figure3. Average score in the three positioning tasks (on the left) and number of subjects, who correctly reported the spoon position (on the right) across the four viewing modes with animation.

When animation is added the average score and the number of subjects who correctly perceived the position varies across viewing modes (Figure 3). These perceptual differences across viewing modes were found to be significant (p<0.001). In general, if we do not take into account the relative position of the spoon to other objects in the scene (positioning task 1, 2 or 3), the order of displays from the best to the worst score for the dynamic positioning task is: D4 = D2 > D3 = D1, where the perception at D2 and D4 are statistically same (p-value=0.999). The perception at D2 is significantly better than at D3 and D1 (p-values=0.002 and < 0.001). The perception at D4 is significantly better than at D3 and D1 (p-value=0.001 and <0.001). The perception at D3 and D1 is same (p-values=0.411), which means that by adding animation to the display mode, users achieved comparable results as with the head tracking mode.

However, Figure 3 suggests that the effect of the viewing mode on subjects' perception of position in dynamic tasks depends on the relative position of the spoon. Indeed, the interaction between the relative position of the spoon (positioning task 1, 2 or 3) and the viewing mode is significant (p-value=0.002). However, the position perception differs across the three tasks only in display mode D3 with the task 2 being the best perceived (p-values = 0.001). Further, in D1, D2 and D4 displays there is no difference across the three tasks (p-values=0.634, 0.603, 0.077). Therefore we conclude that at the positioning task 2, i.e. when the spoon is relatively close to the bowl, the additional use of head tracking improves the perception to a comparable level as is achieved with stereo vision or stereo with head tracking viewing mode.

Further, we directly compared the scores from the static and dynamic positioning task via MIXED analysis on the difference of the scores. On average across all 4 viewing modes and 3 positioning tasks the 36 subjects achieved better results in the perception of the position when animation was added, but these improvements were found significant in monoscopic vision only (p-value<0.05).

7 Conclusion

This paper has presented an experiment designed to investigate the influence of the head tracking on the depth and distance perception in a virtual environment as well as the impact of the animation on our judgment. Usual navigation task in a virtual environment, including games, lies in the movement and rotation of the camera representing a user. The user thus virtually walks or runs and the environment is moving and changing around him/her. Interaction is provided by picking up the selected object or pointing at it. In an environment targeted for a feeding simulation, this interaction is completely different in terms that the user is not moving but sitting in front of the child, the environment is almost static and their interaction with the baby model is based on the navigation of the spoon into its mouth and assessment of the baby's reactions and behaviour. Therefore, the perception of objects and the entire area depends mainly on the movement of the child and right combination of depth cues.

According to the participants' questionnaires, they experienced the strongest sense of depth and distance with combination of stereo vision with head tracking (1.92 in average, where 1 is the strongest sense). Both these viewing modes had individually similar rating (2.39 for stereo vision and 2.56 for head tracking). The monoscopic vision had the worst rating of 3.5.

There is evidence of effectiveness of stereoscopic viewing, which confirms previous studies, for example [12, 13, 15]. Surprisingly, the influence of the head tracking is not as strong with all the tasks as we expected. This might be caused by the fact that participants needed to be encouraged to move towards the baby model several times as they tended to forget.

7 Future Work

In terms of developing more efficient interactive VR environment, more research is needed to assess the effect of the animation on the depth perception and interaction with the objects in 3D space. Therefore, in our future work, we will include more reactions and behaviour related animations of the baby model and investigate how the application can benefit from them. Casting shadows will be added to increase the accuracy of object positioning as they are widely recognized as strong depth cues. In addition, ideally, LCD shutter glasses should be used for better colour perception and essential experience because it is well known that prolonged use of red-cyan glasses can cause headache and strained eyes. To make a user's movement more natural, head tracking should be provided without restriction by devices, for example using web cameras, which are now built-in in most computers.

In future we also plan to extend our experiment to study how other audio and visual components will influence a user's interaction and immersion in the application.

Acknowledgement

This research is funded by EPSRC/Warwick Innovative Manufacturing Research Centre (WIMRC), The University of Warwick, Coventry, UK, projects numbers: R.ESWM.9221, R.PSAA.9221.The authors would like to thank Mrs Debbie Castle, the manager in the University of Warwick nursery, for interesting feeding discussions and all people who participated in the experiment.

References

[1.] RIVA, G., BOTELLA, C., LÉGERON, P., OPTALE, G.: *Cybertherapy: Internet and Virtual Reality as Assessment and Rehabilitation Tools for Clinical Psychology and Neuroscience.* Amsterdam: Ios Press, 2004.

[2.] RIVA, G., WIEDERHOLD, B. K., MOLINARI, E.: *Virtual Environments in Clinical Psychology and Neuroscience.* Amsterdam: Ios Press, 1998.

[3.] COOPER, P., STEIN, A.: *Childhood feeding problems and Adolescent eating disorders*. London and New York: Routledge, 2006.

[4.] SHERMAN, W. R., CRAIG, A. B.: *Understanding Virtual Reality, Elsevier Science*, USA, 2003.

[5.] PETRASOVA, A., FARRER, J. V., CZANNER, S., CHALMERS, A., WOLKE, D.: *User Interface for Assisting Babies with Feeding Disorders*, 2009 International Conference on CyberWorlds, IEEE Computer Society, CA, USA, 2009.

[6.] PLUMERT, J. M., KEARNEY, J. K., CREMER, J. F.: *Distance perception in real and virtual environments,* In Proceedings of Applied Perception in Graphics and Visualization, ACM, New York, USA, pages 27-34, 2004.

[7.] BODENHEIMER, B., MENG, J., WU, H., NARASIMHAM, G., RUMP, B., MCNAMARA, T. P., CARR, T. H., RIESER, J. J.: *Distance estimation in virtual and real environments using bisection*, Applied Perception in Graphics and Visualization, ACM, New York, USA, pages 35-40, 2007.

[8.] WILLIAMS, B., JOHNSON, D., SHORES, L., NARASIMHAM, G.: *Distance Perception in Virtual Environments*, Applied Perception in Graphics and Visualization, Los Angeles, California, ACM New York, USA, 2008.

[9.] GOOCH, A. A., WILLEMSEN, P.: Evaluating *Space Perception in NPR Immersive Environments*, In Proceedings of the 2nd international symposium on Non-photorealistic animation and rendering, ACM, New York, USA, pages 105-110, 2002.

[10.] JACOBSON, L.: *Cognition, Perception and Experience in the Virtual Environment: Do you see What I See?*, International conference on Computer Graphics and Interactive Techniques, ACM New York, USA, 1996.

[11.] BIGOIN, N., PORTE, J., KARTIKO, I., KAVAKLI, M.: *Effects of Depth Cues on Simulator Sickness,* In proceedings of the first international conference on immersive telecommunications, ICST, Brussels, Belgium, 2007.

[12.] HU, H. H., GOOCH, A. A., THOMPSON, W. B., SMITS, B. E., RIESER, J. J., SHIRLEY, P.: *Visual cues for imminent object contact in realistic virtual environment,* Proceedings of the conference on Visualization, IEEE Computer Society Press, CA, USA, 2000.

[13.] LO, C. H., CHALMERS, A.: *Stereo Vision for Computer Graphics: The Effect that Stereo Vision has on Human Judgements of Visual Realism*, In Proceedings of Spring Conference on Computer Graphics, ACM, New York, USA, pages 109-117, 2003.

[14.] LO, C. H., DEBATTISTA, K., CHALMERS, A.: *Selective Rendering for Efficient Ray Traced Stereoscopic Images*, The Visual Computer, Springer Berlin / Heidelberg, 2009.

[15.] HUBONA, G. S., WHEELER, P. N., SHIRAH, G. W., BRANDT, M.: *The relative Contributions of Stereo, Lighting, and Background Scenes in Promoting 3D Depth Visualization*, ACM Transactions on Computer - Human Interaction, ACM, New York, USA, pages 214 - 242, 1999.

[16.] MATHER, G.: *Foundations of Perception*, Psychology Press, Taylor&Francis group, Hove and New York, pages 270-296, 2006.

[17.] MENDIBURU, B.: *3D movie making, Stereoscopic Digital Cinema from Script to Screen*, Focal Press, USA, 2009.

[18.] MURRAY, J.: *Some Perspectives on Visual Depth Perception*, ACM SIGGRAPH Computer Graphics, ACM New York, USA, 1994.

[19.] JONES, J. A., SWAN, J. E., SINGH, G., KOLSTAD, E., ELLIS, S. R.: *The effects of Virtual Reality, Augmented Reality, and Motion Parallax on Egocentric Depth Perception*, In

Proceedings of Applied perception in graphics and visualization, ACM, New York, USA, pages 9-14, 2008.
[20.] SKALA, T., TODOROVAC, M.: *Comparison of stereo displaying techniques in POV-Ray 3D generated scenes*, In Proceedings of iiWAS2008, Linz, Austria, 2008.
[21.] MESSING, R., DURGIN, F. H.: *Distance Perception and the visual Horizon in Head-Mounted Displays*, ACM Transactions on Applied Perception, ACM, New York, USA, pages 234-250, 2005.
[22.] JMonkey game engine. Online Document, September 2009. http://www.jmonkeyengine.com/
[23.] LEE, J. C.: *Head tracking for desktop VR displays using Wii remote*, Online document, September 2009. http://johnnylee.net/projects/wii/
[24.] MASTOROPOULOU, G., DEBATTISTA, K., CHALMERS, A., TROSCIANKO, T.: *The influence of sound effects on the perceived smoothness of rendered animations.* In APGV' 05: Proceedings of the 2nd symposium on Applied perception in graphics and visualization, ACM, New York, NY, USA, pages 9-15, 2005.

Current address

Alena Petrasova
Warwick Manufacturing Group
International Digital Laboratory
The University of Warwick
Coventry, CV4 7AL
United Kingdom
Tel. : +44 (0) 24 765 74587
e-mail: a.petrasova@warwick.ac.uk

Silvester Czanner, RNDr., PhD
Warwick Manufacturing Group
International Digital Laboratory
The University of Warwick
Coventry, CV4 7AL
United Kingdom
Tel. : +44 (0) 24 765 28028
e-mail: s.czanner@warwick.ac.uk