

Moving Towards Consistent Depth Perception in Stereoscopic Projection-based Augmented Reality

S. Schmidt¹, G. Bruder² and F. Steinicke¹

¹Universität Hamburg, Human-Computer Interaction, Germany

²University of Central Florida, USA

Abstract

Stereoscopic projection-based augmented reality (AR) is a promising technology for creating an effective illusion of virtual and real objects coexisting within the same space. By using projection technology, two-dimensional (2D) textures as well as three-dimensional (3D) virtual objects can be displayed on arbitrary physical objects. However, depending on the geometry of the projection surface, even a single virtual object could be projected with varying depths, orientations, and forms. For these reasons, it is an open question whether or not a geometrically-correct projection leads to a consistent depth perception of the AR environment.

We performed an experiment to analyze how humans perceive depths of objects that are stereoscopically projected at different surfaces in a projection-based AR environment. In a perceptual matching task the participants had to adjust the depth of one of two visual stimuli, which were displayed at different depths with varying parallaxes, until they estimated the depths of both stimuli to match. The results indicate that the effect of parallax on the estimation of matching depths significantly depends on the participant's experience with stereoscopic display. Regular users were able to match the depths of both stimuli with a mean absolute error of less than one centimeter, whereas less experienced users made errors in the range of more than 2cm on average. We performed a confirmatory study to verify our findings with more ecologically valid projection-based AR stimuli.

CCS Concepts

•Human-centered computing → Mixed / augmented reality; Empirical studies in HCI;

1. Introduction

Generating a spatial environment that allows a consistent coexistence of real as well as virtual objects is one of the major purposes of AR; however, it still imposes one of the major research challenges of AR. An accurate perception of spatial relationships between virtual and real-world objects is a crucial factor in this context. Traditional AR systems such as head-mounted displays (HMDs) and handheld devices are limited in their capability of providing a consistent spatial impression as various studies showed in the past [KSF10]. Novel near-eye displays attempt to overcome this limitation, for example by using light fields. A different approach is to detach the display technology from the user by projecting virtual information directly onto the surface of realworld objects. In the recent past, this *projection-based AR*, sometimes also referred to as *3D projection mapping* or *spatial augmented reality (SAR)*, was used in a variety of applications, e. g. for simulating different materials of a single object or to transform a whole room into an interactive display (e. g., [JBOW15, JSM* 14]). In most of these applications, a 2D texture is projected onto a 3D surface. However, in particular the use of stereoscopic three-dimensional (S3D) display allows to present an additional dimension in the projection-based AR setup. Using S3D, three-dimensional virtual objects as

well as 2D textures are projected onto real-world 3D geometry, for instance, physical block models of buildings in architecture or exhibits in interactive museums. Such 3D real-world geometry usually consists of several surfaces providing different depths, orientations or forms. Hence, depending on the user's viewpoint, a stereoscopically displayed object would be projected onto different surfaces. Cue conflicts, which arise due to different egocentric distances and stereoscopic parallaxes when different projection surfaces are involved, might cause perceptual differences at the edges of real-world surfaces [RVH13].

It is a challenging question how such conflicts affect the spatial perception of stereoscopically presented 3D objects. Furthermore, so far it is unknown if those visual conflicts could be reduced by perceptually-adapted projections, which compensate how objects are projected onto the surface; even if the perceptually-adapted manipulations lead to geometrically incorrect projections. To investigate these perceptual differences, we performed an experiment in which we analyze the effects of stereoscopic parallax on human perception of consistent depth when stimuli are projected onto different surfaces. The results provide important insights in how depth is perceived in stereoscopic projection-based AR setups between different user groups.

2. Related Work

In this section, we summarize previous work in the field of projection onto real-world surfaces, vergence-accommodation conflicts as well as depth perception in virtual environments (VEs).

2.1. Stereoscopic Projection-based SAR

At the end of the last century, Raskar et al. [RCWS98] demonstrated a prototypic implementation of projection-based AR by registering a virtual 3D model with the underlying 3D physical object in order to overlay additional virtual content. Since then, several projection-based AR setups have been introduced and revised, for example, *Shader Lamps* [RWLB01], *Office of the Future* [RWC*98] and *Emancipated Pixels* [UUI99]. In most of these setups 2D textures were projected onto a 3D geometry, i. e., the virtual information is displayed in monoscopic 2D onto the physical surface. More recently, Jones, Benko and Wilson introduced the *RoomAlive* [JSM*14], *IllumiRoom* [JBOW15] and *Mano-a-Mano* [BWZ14] setups, which allow to monoscopically display virtual objects at any arbitrary 3D location. However, most of the mentioned setups do not provide stereoscopic display, but rather rely on monoscopic cues such as view-dependent perspective to convey the sense of a spatial presence of the virtual object. Using stereoscopic display allows to project virtual 3D objects onto the real-world 3D geometry. In this context, objects may be displayed with *negative*, *zero*, or *positive* parallax, corresponding to appear in front, at, or behind the surface. In the case of *zero parallax* objects appear on the projection surface and can be naturally viewed, i. e., the eyes focus and converge to the same points on the surface. In contrast, objects that appear in front of or behind the projection surface usually result in vergence-accommodation conflicts described in the next section.

2.2. Vergence-Accommodation Conflict

In a natural viewing situation, the vergence stimulus and focal stimulus are at the same distance and therefore the *vergence distance*, i. e., distance to the object to which the eyes converge, and the *focal distance*, i. e., distance to the object at which the eyes focus to sharpen the retinal image, are consistent with each other. However, when viewing a projection-based AR scene with stereoscopic display, the focal distance is fixed at the distance from the eyes to the surface at which the two images for left and right eye are projected, whereas the vergence distance differs depending on the position in space where the object is simulated. This discrepancy results in the well-known *vergence-accommodation conflict* [Pal99]. In order to see an object sharply without double vision, the human viewer must counteract the neural coupling between vergence and accommodation by accommodating to a different distance than the distance to which the eyes converge. Unfortunately, this vergence-accommodation conflicts may result in visual fatigue, visual discomfort and spatial misperception as previous work has shown [HGAB08]. In particular, several studies reported a tendency towards depth underestimation in VEs where virtual objects are displayed with positive parallax as in head-worn AR and HMD as well as projection-based VR (for a review see [RVH13]). Furthermore, Bruder et al. [BSOL15] reproduced this effect in a large

ten meter projection system and also revealed a distance overestimation for close objects at negative parallaxes. Nevertheless, the influence of different technical and human factors on depth perception in VEs is still object of investigation.

2.3. Depth Perception in AR

Many individual depth cues are used by the visual system to provide an estimate of relative and absolute distances of objects within an image. According to Cutting and Vishton [CV95], the importance of different depth cues depends on the distance of the viewer to the object of interest. In the personal space, binocular disparity as well as accommodation provide the most accurate depth cues. In this context, Ellis and Menges [EM98] conducted experiments to investigate the effects of stereoscopic vision on depth perception in AR using HMDs. Their results show a main effect of stereoscopy on depth estimation in the near visual field, resulting in greater accuracy. Furthermore, Broecker et al. [BST14] investigated how different cues affect the depth perception in a view-dependent near-field projection-based AR setup. Their results suggest that head-tracking improves depth perception though this trend was not significant. Furthermore, their setup did not consider stereoscopic display at all. Benko et al. [BJW12] performed a depth perception study in the *MirageTable* setup. They found that participants were reasonably accurate in their estimates, with an average depth estimation error of ca. 2.4cm. However, the focus of their work was on the technical realization, and vergence-accommodation conflicts and the effects of stereoscopic parallax were not further analyzed.

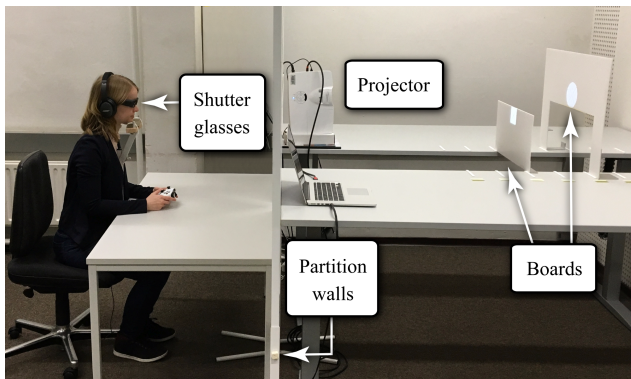
Overall, the results of these studies suggest that both binocular disparity and accommodation provide important depth information in projection-based AR and therefore can be used to resolve ambiguities created by other perceptual cues. So far, the interaction with other depth cues such as brightness differences or blur, which are inevitable when virtual content is projected onto different depth planes via one single projector, is mostly unknown. Schmidt et al. [SBS16] investigated the manipulation of perceived depth of real-world objects by using visual illusions based on projecting different color, luminance contrast or blur effects onto their surfaces. The results of the study suggest that perceived depth of objects can be altered, although binocular vision dominated the other cues in all tested conditions.

3. Psychophysical Experiment

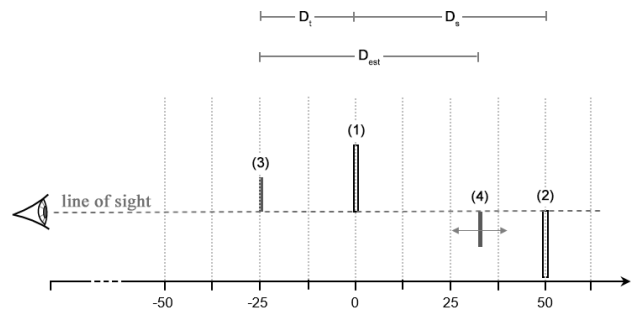
In this section we describe the experiment that was conducted to analyze effects of stereoscopic parallax on depth perception in projection-based AR environments. The experiment involved a perceptual matching task, in which participants were shown two connected visual stimuli top-bottom at different depths. Their task was to adjust the depth of one of the stimuli until they estimated that depths of both stimuli matched.

3.1. Participants

A total of 20 participants, 18 male and 2 female (aged from 20 to 38, $M=28.2$) were recruited through advertisements. All participants were students or members of the local department of computer science. All participants had normal or corrected-to-normal



(a)



(b)

Figure 1: Illustration of (a) the experimental setup and (b) the analyzed distances. In the shown configuration the upper projection surface (1) is stationary. The lower projection surface (2) and the target object (3) are placed at one of the illustrated distance marks at the beginning of each trial, while the object (4) was controlled by the user.

vision; six wore glasses during the study. One participant suffers from a mild kind of astigmatism. No other vision disorders, such as color or night blindness, dyschromatopsia, or a strongly impaired eyesight, were reported.

3.2. Materials

The experimental setup is illustrated in Figure 1. During the experiment, participants were facing a real scene, which was augmented with two virtual objects via 3D projection mapping. To display the stereoscopic imagery we used an active shutter 3D system, including a 3D-capable Optoma HD20 projector as well as compatible RF shutter glasses. The projector was placed out of view behind two partition walls that also hid the mounting of the projection surfaces (cf. Figure 1). Through the restriction of their view, participants were restrained from applying a strategy such as using reference points for their depth estimation. In order to keep the projection center and the participant's eyes vertically aligned throughout all trials, a chin rest was used to fix the participant's head position. The virtual scene was projected onto two planar foam boards with a smooth, white-colored, diffusely reflecting surface. The boards were placed one above the other, both facing the participant. The initial alignment of the projector and the boards was performed in a one-time calibration step that utilized the RoomAlive framework [JSM^{*} 14] as well as a custom marker tracking implementation. During the main experiment, the boards were shifted manually using stopper at pre-defined positions.

Since the boards were shifted in depth during the experiment, we aligned the focal plane of the projector with the board at medium distance. Hence, all objects projected onto this board appeared sharp, whereas objects projected onto the other board were slightly out of focus depending on the current surface distance. Although this blur effect was barely noticeable with naked eye, a possible correlation between defocus and the estimated distance is considered in the analysis. The difference of illuminance between an object projected at maximum distance and one projected at minimum distance was 400lx.

For the visual stimuli we used a flat textured square and a circle with a randomized size between 8cm and 12cm. We chose these reduced-cue stimuli in order to focus on binocular disparity and accommodation as the main available depth cues. The different shapes were used because of a pre-test, which revealed that participants heavily focused on the edges instead of the objects' depths when two squares had to be matched. These pre-tests also showed that a texture helped the participants to focus on the virtual objects, which is also in line with existing literature on surface textures rendered with stereoscopic displays [TGCH02]. Furthermore, in our first experimental setup we dynamically adapted the size of the virtual objects in such a way that the retinal size was kept constant regardless of their current distance. However, this was rated as unnatural by multiple testers and therefore was discarded in the current setup.

During the experiment, the participant was required to control one of the projected virtual objects along its z -axis via a connected gamepad. The movement was not limited to a minimum or maximum value. The smallest achievable change of position of the virtual object with the gamepad was 1mm.

3.3. Design and Procedure

Prior to the study, the interpupillary distance (IPD) of every participant was measured to provide a correct stereoscopic rendering of the virtual scene in the trials. We verified each participant's ability to perceive binocular depth with the Titmus test, followed by a graded circle test to evaluate their stereoscopic acuity [FS97]. After passing all pre-tests, participants were instructed to sit in an upright position as explained in Section 3.2. They received detailed instructions on how to perform the required task. To familiarize with the setup and the stereoscopic stimuli, every participant passed an initial training phase before the actual experiment started. These trials were excluded from the analysis.

For the main part of the study we followed a mixed factorial design with the three independent variables *surface offset*, *target offset* and *moving board*. We define the *surface offset* as the relative

distance from the movable board to the stationary board with a positive offset indicating that the movable board is further away than the stationary one as seen from the participant. Similarly, we define the *target offset* as the relative distance from the controllable object to the target with a positive offset indicating that the controllable object is further away than target as seen from the participant. For every participant, the moving board was chosen randomly in the beginning of the experiment and afterwards kept constant during the entire session. Overall, the decision if the upper or the lower board was moving was counterbalanced between all participants.

In order to investigate possible correlations between depth estimation and the offset between projection surfaces as well as the relative position of the virtual stimuli compared to the surfaces, we used two different configurations. In both of them the stationary board was placed at a fixed egocentric distance of 200cm in front of the participant. In the first configuration, the target was always projected onto the stationary board with zero-parallax. The dynamic board was moved between nine pre-defined positions with the following relative offsets to the stationary board: $D_s \in \{-50, -37.5, -25, -12.5, 0, 12.5, 25, 37.5, 50\}$ cm (cp. Figure 1b). In the second configuration, the stationary board was still placed at its initial position while the second board was positioned at a relative offset of 50cm behind the first one. Furthermore, the target was moved between five different locations with offsets $D_t \in \{-25, 0, 25, 50, 75\}$ cm. This corresponds to a relative positioning of the target in front of both boards, at the same depth as the first board, between the boards, on the same depth as the rear board and behind both boards, respectively.

According to the current condition, both the movable board and the target were moved to one of the pre-defined distances D_s and D_t before each trial. The shape of the target (circle or square) as well as its size were chosen randomly. The second virtual object, which was controlled by the participant, was initialized at a random offset in the range of -60 cm to 60 cm. For each trial, participants had to move the controllable object along its z -axis to the perceived depth of the corresponding stationary target with the gamepad. Since the target and the controllable object differed in their size, participants had to rely on their depth perception instead of matching the objects' edges. The relative distance between the estimated depth of the controllable object and the target position was recorded as the dependent variable of the study.

In order to restrain participants from comparing the target position between trials, all conditions of both configurations were presented fully randomized. In addition, we introduced two trials at the beginning of each experiment session (after the training phase) in order to verify the participant's ability to perform the task correctly. In these two verification trials, both boards were placed at the same depth and the target was projected 25cm in front of or behind the boards, respectively. Thus, the task was reduced to adjust the same parallax for both visual stimuli and therefore should be solvable for every person with normal stereoscopic vision and correct understanding of the task. Including the verification trials, we tested 15 different conditions. This is because one condition was equivalent in the first and second configuration ($D_s = 50$ cm and $D_t = 0$ cm) and therefore was only included once. After presenting every condition, they were repeated a second time, again in ran-

domized order. This results in an overall number of 30 trials per participant. Between two trials, the participants had to close their eyes, which was signaled to them via headphones. The headphones also used active noise cancellation in order to minimize the bias through noise caused by the repositioning of the boards.

After the study, the participants completed a questionnaire to provide qualitative feedback as well as some demographic information. The total time per participant including questionnaires, instructions, training, experiment and debriefing was around half an hour.

Considering previous results in the literature and the depth cues described in Section 2 our hypotheses are:

- H1** Increasing the surface offset leads to increased absolute errors in matching depths estimates.
- H2** Increasing the target offset leads to increased absolute errors in matching depths estimates.
- H3** Participants experienced in the usage of S3D glasses provide more accurate estimations for matching depths.

Although other, sometimes contrary cues also affect the depth perception as mentioned in Section 2.3, we still expect convergence and accommodation to be the most dominant cues in near-field AR, resulting in an underestimation of the distance to objects exhibiting positive parallax and overestimation of the distance to objects exhibiting negative parallax. The third hypothesis is mainly based on observations made in previous experiments. Participants who experienced stereoscopic display only very occasionally often reported difficulties in judging distances or focusing on virtual 3D objects, especially when these objects exhibited a strong parallax.

3.4. Results

For analyzing the results of the psychophysical experiment we discarded two participants from the data, since their estimated depth extremely deviated from the target depth in the verification trials. Besides, five data points with values more than three times the interquartile range were considered as extreme outliers and therefore also excluded from the analysis. On the resulting data set we conducted multiple JZS Bayes factor ANOVAs [RMSP12]. Over the last years it has become increasingly apparent that the Bayesian approach to data analysis comes with considerable advantages over classical statistics, both theoretical and practical (e.g. [Die11]; for a systematic overview of more than 1500 articles reporting Bayesian analyses in psychology see [vdSWR*17]).

3.4.1. Surface Offset

Figure 2(a) shows the mean absolute differences between estimated distances D_{est} and target distance D_t for the experiment conditions in which the target object was always presented at an egocentric distance of 200cm, i. e., the target distance always matched the distance of the stationary board. The surface offset on the x -axis indicates the relative distance from the movable board to the stationary board with a positive offset indicating that the movable board is farther away than the stationary one as seen from the participant. Since the target object was always presented at a distance of 200cm, it was thus presented with exactly zero parallax on the stationary surface, whereas the controllable object was projected with parallax

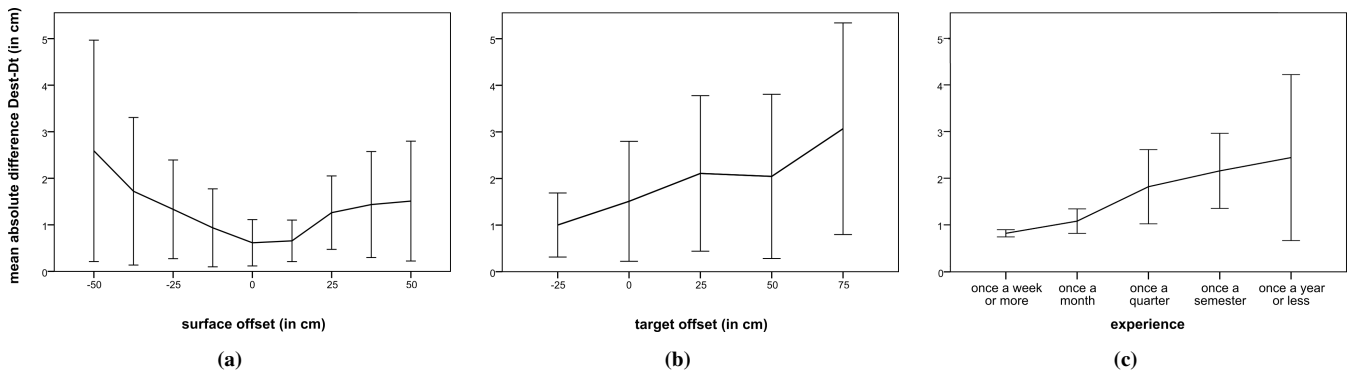


Figure 2: Results of the first experiment: Pooled differences between the estimated distance and target distance ($D_{est} - D_t$) (a) for surface offsets D_s , (b) for target offsets D_t , and (c) for the experience of participants with S3D glasses measured on a 5-point Likert scale.

in case its depth matched the depth of the target. The vertical bars show the standard deviation of the mean.

A JZS Bayes factor ANOVA with default prior scales revealed that $H1$ was preferred to the null model by a Bayes factor of $B_{10} = 3353.142$. Therefore, the data provide very strong evidence for the hypothesis $H1$ that larger surface offsets lead to increased absolute errors in matching depths estimates. However, considering the results of every participant separately, we observed individual trends. Participants, who adjusted the object in front of the target for negative surface offsets, moved the object behind the target for positive surface offsets and vice versa. While the tendency towards underestimation of the distance to target objects that are displayed with positive parallax can be explained by the vergence-accommodation conflict, the opposite trend has to be induced by depth cues other than accommodation. We evaluated a correlation of the moving board and the reported strategies with the individual trends of participants, but could not find a reportable effect.

3.4.2. Target Offset

Figure 2(b) shows the mean absolute differences $D_{est} - D_t$ pooled over target offsets D_t . The vertical bars show the standard deviation of the mean. According to our hypothesis $H2$ we expected increased absolute errors in matching depths estimates with increasing target offsets D_t . In order to evaluate this, we performed another Bayes factor ANOVA with default prior scales, resulting in a Bayes factor B_{20} of 1501.625. According to Raftery [Raf95] this corresponds to a very strong evidence against the null model in favor of $H2$. Furthermore, we expected an underestimation of depth at all target offsets due to the vergence-accommodation conflict. Although this trend can be observed for a subgroup of participants, we also registered an opposite trend as in the previous configuration considering the surface offsets. In general, we observed a higher standard deviation for target positions further away from the user, which could indicate a dependency of the estimates from egocentric distance rather than target offset.

3.4.3. Experience

For measuring the experience of participants with S3D glasses, we used a 5-point Likert scale with values 'once a week or more',

'once a month', 'once a quarter', 'once a semester' and 'once a year or less'. Each option was chosen by 2 to 5 participants. For every participant, we averaged the means of absolute error in the estimated distance of the 13 different conditions. The results are plotted in Figure 2c. For the 18 participants of the experiment a trend can be observed, suggesting a higher accuracy of distance estimation with increasing experience with S3D glasses. In order to validate this assumption, we formed two subgroups of regular users, who reported to wear S3D glasses at least once a month, and occasional users. A two-sample JZS Bayes t-test with default prior scales [RSS*09] resulted in a Bayes factor of 1.799, suggesting a weak evidence in favor of $H3$ against the null model. To clarify this result, we analyzed the experiment data with an additional t-test, which revealed a significant difference between regular users ($M=0.98, SD=0.24$) and occasional users ($M = 2.16, SD = 1.20$); $t(16) = 2.15, p = .047$. This further supports our hypothesis $H3$; however, a larger sample could be considered in future experiments in order to allow a more differentiated analysis for several levels of experience.

3.5. Discussion

Overall, we observed large variance in the responses, which increased with larger offsets between the projection surfaces, whereas depth estimation was more accurate for small surface offsets. However, the experiment revealed different trends that are not correlating with the strategies, which were reported in the post-questionnaire. To verify this observation, three participants repeated a shortened version of the experiment, in which they had to wear an additional Pupil Labs headset for binocular eyetracking. An analysis of the sample eye tracking data did not reveal a dependency between the focused board and the observed trends. Participants moved the controllable object back and forth until it leveled off at a depth, they perceived as correct. In particular during the fine tuning at the end of each trial, their gaze switched between the visual stimuli several times.

While different strategies to solve the task do not seem to have an impact on the estimation error, the results of the experiment indicate a correlation between the experience of participants with S3D glasses and absolute difference between estimated depth and target

depth. In particular, for an increasing surface offset the measured absolute error turned out to be higher for unexperienced participants. Including the finding that distance estimation was nearly veridical for experienced participants, although they were confronted with the same vergence-accommodation conflict, another interpretation is admitted that considers additional depth cues to binocular disparity. Participants with less experience in wearing S3D glasses could have difficulty focusing the two visual stimuli with increasing difference of the parallaxes and could therefore make use of other cues, consciously or unconsciously. This would be in line with the qualitative feedback that was provided in the questionnaires. The integration of luminance contrast between both visual stimuli could cause an underestimation of the distance to objects exhibiting positive parallax and an overestimation of the distance to objects exhibiting negative parallax as observed for a subgroup of participants. The opposite results may be caused by a subliminal depth compression. A common technique in 3D filmmaking is to reduce depth differences of the scene in order to minimize visual discomfort of the viewers [LHW*10]. Regarding the conducted experiment, the parallax difference between controllable object and target could be reduced by moving the object closer to the projection surface instead of further away as expected due to the vergence-accommodation conflict. A smaller parallax difference results in a more comfortable viewing experience and could therefore influence the participant's depth estimation.

Considering the overall results of the first experiment, the question arises, whether a perceptual motivated correction of object depths improves spatial perception of a projection-based AR environment or if no perceptual inconsistencies occur for geometrically correct projections. For further investigation of this question, we decided to perform a follow-up study described in the next section.

4. Confirmatory Study

The results of the first psychophysical experiment suggest a correlation of the distance between both projection surfaces and the depth estimation error, which strongly depends on participant's experience with S3D. We conducted a confirmatory study in order to test whether a compensation of this error results in a perceivable improvement of the spatial impression in near-field projection-based AR. Using the setup described in Section 3.2 the participants saw two virtual 3D objects; one was displayed without any modifications whereas the other's halves were shifted against each other according to the depth estimate error of the participant in the first experiment. The participants then performed a 2AFCT, deciding for which of the two presented objects they had a more consistent spatial impression.

4.1. Participants

From the set of participants of the first experiment, we recruited 12 participants, 11 male and 1 female (aged from 20 to 38, $M=29.25$). The sample equally represented the three different behaviors, which were identified in the first psychophysical experiment.

4.2. Design and Procedure

In the confirmatory study we followed a repeated measures within-subjects design, which involved the surface offset ($D_s \in \{-50, -37.5, -25, -12.5, 0, 12.5, 25, 37.5, 50\}$ cm), target offset ($D_t \in \{-25, 0, 25, 50, 75\}$ cm) and moving board (upper/lower) as independent variables. Possible combinations of surface distance and target distance were the same as in the first experiment. However, this time each participant performed trials both with the upper and the lower board moving. In order to reduce the time for changing the boards' positions between the trials, all conditions were grouped according to the moving board. Therefore, the moving board only switched once after the participant finished all conditions of the first, randomly chosen group.

In order to simulate a realistic projection scenario, we used a virtual textured 3D model of a skyscraper as visual stimuli for the confirmatory study. For a better comparability, the size of both skyscrapers was kept constant through the experiment. As described in Section 3.3 both projection surfaces as well as the virtual objects were positioned according to one of 13 possible configurations before each trial. Unlike the first experiment, the upper and lower parts of both objects were not positioned independently of one another. Instead, they were placed either exactly one above the other or with a slight depth shift as described before. Each configuration was repeated twice with the geometric correct skyscraper projected on the left and the right, respectively.

In summary, participants performed 13 conditions with 2×2 repetitions each, resulting in 52 presented trials, which were randomly presented. Overall, one session took around 20 minutes to complete.

4.3. Results

For analyzing the results of the 2AFCT we ran one-sample JZS Bayes factor t-tests with default prior scales [RSS*09] and a null value of the mean of 50 for each level of D_t and D_s .

To investigate whether the participants perceived a qualitative difference between the perceptually adapted and the geometrically correct projection, we compared the following two models:

M0 Random decision.

M1 Non-random decision.

The resulting Bayes factors are listed in Table 1. For different levels of the target offset D_t the t-tests resulted in Bayes factors B_{10} ranging from 0.287 to 0.378. According to Raftery [Raf95] this corresponds to a positive evidence of the hypothesis that participants were guessing randomly in case the target was positioned at $D_t \in \{-25, 0, 25, 50\}$ cm and only a weak evidence when $D_t = 75$ cm.

For surface offsets D_s the t-tests revealed Bayes factors B_{10} between 0.298 and 0.915, suggesting only a weak evidence against M_1 . One exception is the Bayes factor for a surface offset of 0, which is in favor of the alternative model M_1 against the null model M_0 by a factor of about 10.902. However, this result was predictable, since no perceptual adaption of the virtual content should be necessary when both boards are positioned at same depth.

Since there is no strong evidence in favor of either model M_0 or M_1 , we additionally considered the following models:

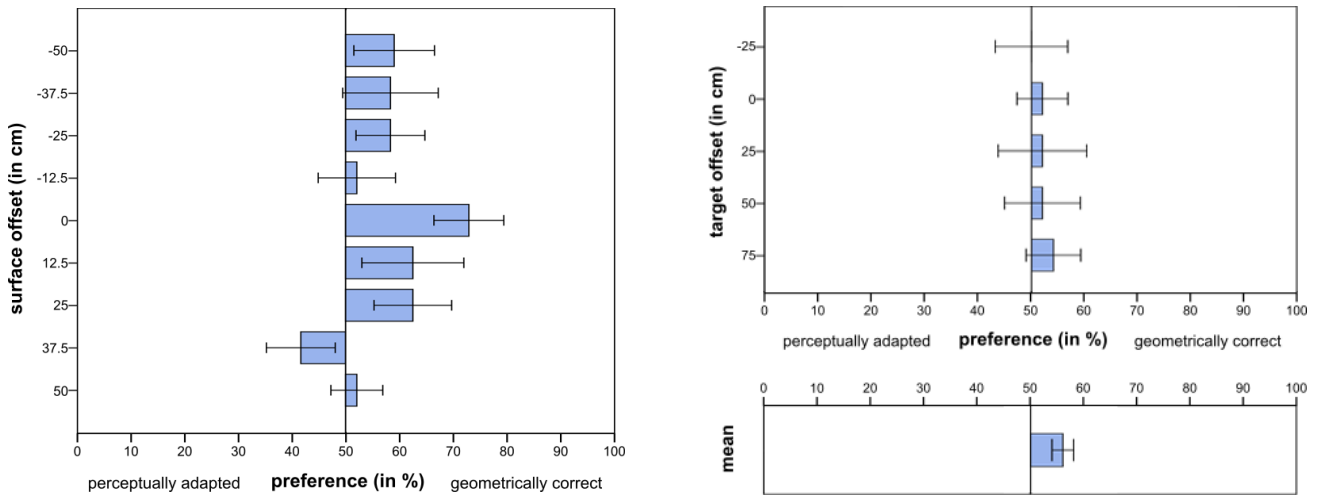


Figure 3: Pooled results of the 2AFCT for the surface offsets D_s (left), the target offsets D_t (right top) and a mean value across all surface and target offsets (right bottom). The x-axis shows a percentage value of how often the geometrically correct projection was judged as more consistent.

Table 1: Bayes factors for comparisons of the models M_1 and M_0 as well as M_3 and M_2 . The first row represents different levels of surface offset D_s (left) and target offset D_t (right).

D_s	-50	-37.5	-25	-12.5	0	12.5	25	37.5	50	D_t	-25	0	25	50	75
B_{10}	0.520	0.416	0.572	0.298	10.903	0.581	0.581	0.572	0.312	B_{10}	0.287	0.312	0.295	0.298	0.378
B_{32}	5.956	3.980	6.913	1.525	235.048	7.097	13.533	0.145	1.876	B_{32}	1.000	1.878	1.435	1.525	3.258

- M2** Preference of perceptually adapted projection.
- M3** Preference of geometrically correct projection.

M_2 and M_3 were tested against one another for varying surface offsets D_s . This allows to investigate the participants’ preferences of either a perceptually-adapted or the geometrically correct projection, assuming a non-random decision. We conducted one-tailed t-tests with a null interval of (0, infinity) and (-infinity, 0), respectively. By dividing the resulting values B_{20} by B_{30} , we got Bayes factors B_{32} as shown in table 1. They suggest that the data favor M_3 over M_2 for six out of seven surface offsets.

For the sake of completeness Bayes factors B_{32} for different target offsets D_t are also listed in table 1, although no strong evidence against one model or the other could be found.

4.4. Discussion

The Bayesian analysis provides indications that perceptually adapted projection is not preferred to geometrically correct projection of the visual stimuli, independent from the individual behavior in the first experiment. Decision rates close to 50% suggest that the 2AFCT was approaching our participants’ sensitivity to differences in depth when using stereoscopic display. However, we also observed a tendency towards the geometrically correct projection, indicating that variances in adjusted distances from the first experiment may be caused by uncertainties and do not reflect a real per-

ceptual difference between the depths of both visual stimuli. It still has to be investigated if the same results can be reproduced in a full-cue environment, when the user’s perception is influenced by other depth cues such as motion parallax.

5. Conclusion

In this paper we presented a psychophysical experiment and a confirmatory study to investigate the effects of stereoscopic parallax on the human depth perception in projection-based AR environments. Such environments typically contain several surfaces with various depths, orientations or forms and, therefore, perceptual differences might occur when virtual objects are stereoscopically projected over multiple surfaces at different depths. In order to evaluate differences in depth perception and consistency of stereoscopically presented depth of virtual objects, we projected visual stimuli at two different surface planes with varying distances to the user. A perceptual matching task was performed, which gives indications on the depth perception in a SAR environment.

First, the results support the hypotheses that increasing offsets between multiple projection surfaces as well as the projection surfaces and projected targets lead to increased absolute errors in estimated depths. However, the relative errors differ between participants and therefore cannot be explained by the vergence-accommodation conflict in each individual case. The observed

trends could be caused by individually perceived and weighted characteristics of a SAR environment such as luminance differences of visual stimuli projected onto different projection surfaces. Considering the variance in the responses, it can be assumed that for most participants estimation of target distances was more difficult for larger surface offsets.

Furthermore, the results indicate that the effect of parallax on the estimation of matching depths strongly depends on the participant's experience with S3D. Participants, who wear S3D glasses at least once a week, were able to match the depths of both stimuli with a mean error of less than one centimeter. This is an interesting result since it suggests that more experienced users perceive VEs displayed in a SAR setup in a different way than less experienced users. However, the confirmatory study revealed a tendency towards preference of a geometrically correct projection of the visual stimuli, independently from the individual behavior of the participants in the first experiment. Considering practical applications of SAR, this could indicate that there is no need for a complex perceptual adaptation of the visual stimuli in order to create a spatially consistent SAR environment. However, it also implies that offsets between physical projection surfaces and stereoscopically projected objects should be reduced to a minimum in order to facilitate perceptual integration of stimuli, in particular for users who are less experienced in the usage of S3D glasses.

Future work should focus on the analysis of the learning curve for reliable depth estimations in S3D environments. Furthermore, we would like to explore full-cue environments, in which the user's perception is influenced by other depth cues such as motion parallax.

References

- [BJW12] BENKO H., JOTA R., WILSON A.: MirageTable: Freehand Interaction on a Projected Augmented Reality Tabletop. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)* (2012), pp. 199–208. 2
- [BSOL15] BRUDER G., SANZ F. A., OLIVIER A.-H., LÉCUYER A.: Distance Estimation in Large Immersive Projection Systems, Revisited. In *Proceedings of the IEEE Virtual Reality Conference (VR)* (2015), pp. 27–32. 2
- [BST14] BROECKER M., SMITH R. T., THOMAS B. H.: Depth Perception in View-Dependent Near-Field Spatial AR. In *Proceedings of the Australasian User Interface Conference (AUIC)* (2014), pp. 87–88. 2
- [BWZ14] BENKO H., WILSON A. D., ZANNIER F.: Dyadic Projected Spatial Augmented Reality. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)* (2014), pp. 645–655. 2
- [CV95] CUTTING J. E., VISHTON P. M.: Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth. *W. Epstein & S. Rogers (Eds.) Vol 5; Perception of Space and Motion* (1995), 69–117. 2
- [Die11] DIENES Z.: Bayesian Versus Orthodox Statistics: Which Side Are You On? *Perspectives on Psychological Science* 6, 3 (2011), 274–290. 4
- [EM98] ELLIS S. R., MENGES B. M.: Localization of Virtual Objects in the Near Visual Field. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 40, 3 (1998), 415–431. 2
- [FS97] FRICKE T. R., SIDEROV J.: Stereopsis, Stereotests, and their Relation to Vision Screening and Clinical Practice. *Clinical and Experimental Optometry* 80, 5 (1997), 165–172. 3
- [HGAB08] HOFFMAN D. M., GIRSHICK A. R., AKELEY K., BANKS M. S.: Vergence–Accommodation Conflicts Hinder Visual Performance and Cause Visual Fatigue. *Journal of vision* 8, 3 (2008), 33–33. 2
- [JBOW15] JONES B., BENKO H., OFEK E., WILSON A.: IllumiRoom: Immersive Experiences Beyond the TV Screen. *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)* 58, 6 (2015), 93–100. 1, 2
- [JSM*14] JONES B., SODHI R., MURDOCK M., MEHRA R., BENKO H., WILSON A., OFEK E., MACINTYRE B., RAGHUVANSHI N., SHAPIRA L.: RoomAlive: Magical Experiences Enabled by Scalable, Adaptive Projector-Camera Units. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)* (2014), pp. 637–644. 1, 2, 3
- [KSF10] KRUIJFF E., SWAN J., FEINER S.: Perceptual Issues in Augmented Reality Revisited. In *Proceedings of the IEEE Symposium on Mixed and Augmented Reality (ISMAR)* (2010), pp. 3–12. 1
- [LHW*10] LANG M., HORNING A., WANG O., POULAKOS S., SMOLIC A., GROSS M.: Nonlinear Disparity Mapping for Stereoscopic 3D. *ACM Transactions on Graphics (TOG)* 29, 4 (2010), 75. 6
- [Pal99] PALMER S.: *Vision Science: Photons to Phenomenology*. A Bradford Book, 1999. 2
- [Raf95] RAFTERY A. E.: Bayesian Model Selection in Social Research. *Sociological Methodology* (1995), 111–163. 5, 6
- [RCWS98] RASKAR R., CUTTS M., WELCH G., STUERZLINGER W.: Efficient Image Generation for Multiprojector and Multisurface Displays. In *Proceedings of the Eurographics Workshop* (1998), pp. 139–144. 2
- [RMSP12] ROUDER J. N., MOREY R. D., SPECKMAN P. L., PROVINCE J. M.: Default Bayes Factors for ANOVA Designs. *Journal of Mathematical Psychology* 56, 5 (2012), 356–374. 4
- [RSS*09] ROUDER J. N., SPECKMAN P. L., SUN D., MOREY R. D., IVERSON G.: Bayesian t Tests for Accepting and Rejecting the Null Hypothesis. *Psychonomic Bulletin & Review* 16, 2 (2009), 225–237. 5, 6
- [RVH13] RENNER R. S., VELICHKOVSKY B. M., HELMERT J. R.: The Perception of Egocentric Distances in Virtual Environments - a Review. *ACM Computing Surveys (CSUR)* 46, 2 (2013), 23. 1, 2
- [RWC*98] RASKAR R., WELCH G., CUTTS M., LAKE A., STESIN L., FUCHS H.: The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays. In *Proceedings of the ACM Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)* (1998), pp. 179–188. 2
- [RWLB01] RASKAR R., WELCH G., LOW K.-L., BANDYOPADHYAY D.: Shader Lamps: Animating Real Objects with Image-Based Illumination. In *Proceedings of the Eurographics Workshop on Rendering Techniques* (2001), pp. 89–102. 2
- [SBS16] SCHMIDT S., BRUDER G., STEINICKE F.: Illusion of Depth in Spatial Augmented Reality. In *Proceedings of the IEEE VR Workshop on Perceptual and Cognitive Issues in AR (PERCAR)* (2016). URL: <http://basilic.informatik.uni-hamburg.de/Publications/2016/SBS16a>. 2
- [TGCH02] THOMAS G., GOLDBERG J. H., CANNON D. J., HILLIS S. L.: Surface Textures Improve the Robustness of Stereoscopic Depth Cues. *Human Factors* 44, 1 (2002), 157–170. 3
- [UIU99] UNDERKOFFLER J., ULLMER B., ISHII H.: Emancipated Pixels: Real-World Graphics in the Luminous Room. In *Proceedings of the ACM Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)* (1999), pp. 385–392. 2
- [vdSWR*17] VAN DE SCHOOT R., WINTER S. D., RYAN O., ZONDERVAN-ZWIJNENBURG M., DEPAOLI S.: A Systematic Review of Bayesian Articles in Psychology: The Last 25 Years. *Psychological Methods* 22, 2 (2017), 217–239. 4