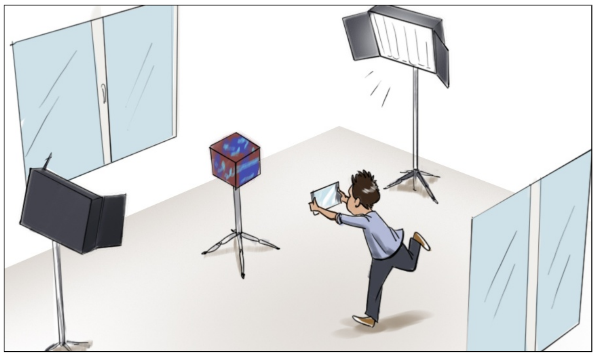
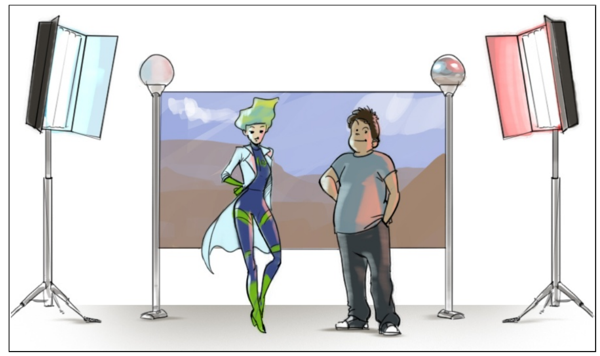
Influence of Animated Reality Mixing Techniques on User Experience

Fabio Zünd[1], Marcel Lancelle[1], Mattia Ryffel[2], Robert W. Sumner[1,2], Kenny Mitchell[2], and Markus Gross[1,2]

[1]ETH Zurich
[2]Disney Research Zurich



(a) *ARTravelers Concept Drawing*



(b) *ARPix Concept Drawing*

Figure 1: *(a): ARTravelers concept drawing, as installed for the blur (Section 4) and latency (Section 5) experiment. The cubical marker in the center as well as the surrounding scene is well lit by the windows and lamps. During the game, the player runs around the cube holding a mobile device and tries to destroy as many virtual targets as possible in a limited time. (b) ARPix concept drawing. A virtual character stands next to an image marker. The user can stand next to it for a group photo, captured by an additional user. The white and silver spheres have a known position and are used to capture the environment lighting. This enables a realistic shading of the virtual character.*

Abstract

We investigate the influence of motion effects in the domain of mobile Augmented Reality (AR) games on user experience and task performance. The work focuses on evaluating responses to a selection of synthesized camera oriented reality mixing techniques for AR, such as motion blur, defocus blur, latency and lighting responsiveness. In our cross section of experiments, we observe that these measures have a significant impact on perceived realism, where aesthetic quality is valued. However, lower latency records the strongest correlation with improved subjective enjoyment, satisfaction, and realism, and objective scoring performance. We conclude that the reality mixing techniques employed are not significant in the overall user experience of a mobile AR game, except where harmonious or convincing blended AR image quality is consciously desired by the participants.

CR Categories: I.2.10 [Artificial Intelligence]: Vision and Scene Understanding—Perceptual Reasoning, I.4.8 [Computer Graphics]: Three-Dimensional Graphics and Realism—, K.8.0 [Personal Computing]: General—Games

Keywords: augmented reality, motion blur, camera motion, realism, computer graphics

1 Introduction

Augmented Reality (AR) features in an increasing number of mobile games and interactive entertainment applications. In visual terms, the presence of real and virtual objects that coexist in the same space [Azuma 1997] is conveyed through the seamless blending of the camera image with rendered computer graphics. Further, the realistic depiction of motion in interactive AR is affected by various reality mixing measures including matched motion blur, reduced latency, and responsiveness to changes in the camera image.

Recently, the perceptual and practical impact on the players' user experience of motion blur in racing games has been studied rigorously with the, perhaps surprising, outcome that the effect is itself not essential to the enjoyment of the game overall [Sharan et al. 2013].

In a virtual reality (VR) scenario, all animated graphical content takes place within the bounds of a display showing the virtual environment, and so the resulting effect of motion blur in this scenario is purely synthetic. However, given that AR brings into play the context of the players' real visual surroundings, we are interested to determine whether the same outcome would be observed. That is, would similar responses arise from a user study examining the effects on enjoyment, sense of realism, satisfaction, and task performance with reality mixing methods.

Real-time computation performance is more challenging for AR than VR in that the balance and blend of rendered graphics must be targeted to extract and process the real camera image to mix and match content seamlessly. Therefore, it is important to determine which reality mixing measures matter for user experience and to what degree must they be accurate. In this article, we reprise the motion blur user study of [Sharan et al. 2013], but for the AR domain and then further investigate the motion effects of varying levels of latency and responsiveness to dynamic environment lighting changes.

2 Related Work

Many studies have analyzed the results of user experiments to assess cognition, perception, task performance and collaboration in AR [Swan(II) and Gabbard 2005]. More recently, a collection of investigations from a visual observation and interpretation standpoint has been surveyed [Kruijff et al. 2010]. Whilst here, we target our investigations upon the influence on the overall user experience for interactive entertainment applications in mobile AR.

The early work by Fischer et al. [Fischer et al. 2006] and a more comprehensive work by Klein and Murray [Klein and Murray 2010] are most relevant for our experiments dealing with compositing camera effects within virtual rendered content. They address the visually most important effects such as vignetting, noise, chromatic aberration, and blur caused by imperfect imaging and camera motion. Like in our experiments, Park et al. [Park et al. 2012] use a video see through AR framework with mobile devices.

Instead of artificially decreasing the quality of virtual content, another approach is to improve the camera image quality. Unfortunately, many of these methods are ill posed and are usually computationally too expensive to run in real time on current mobile devices. Blur kernels in images are often estimated for subsequent deconvolution, i.e. restoration. A number of methods exists that handle only motion blur [Oh and Kim 2014] or defocus blur [Tao et al. 2013]. Others do a joint estimation [Oliveira et al. 2014], but all of these image based techniques are computationally expensive. Other options include estimation of the kernel with an IMU with accelerometers or gyroscopes rigidly attached to the sensor [Joshi et al. 2010], [Bae et al. 2013]. We applied this idea to estimate the required motion blur for the virtual content.

For consistent material environment shading, a mirror sphere visible in the camera image can be used to extract lighting information [Kanbara and Yokoya 2002] in real-time and render materials in a realistic way [Agusanto et al. 2003]. These methods also handle the white balance and color matched illumination to some degree. We follow this approach in our dynamic light environment experiment, except we use a diffuse sphere to effectively sample the optically pre-convolved diffuse materials, which results in a reduced computation reflection mapping for low-powered mobile devices.

Sharan et al. [Sharan et al. 2013] show in experiments that motion blur in a racing game on a video game console is preferred by the users but has no influence on task performance. In a similar way, a psychological study in [Knez and Niedenthal 2008] examined the influence of different in-game lighting conditions to player performance and feelings. They concluded that their participants solved maze levels in a first person shooter game fastest and best in warm (reddish) as compared to cold (blueish) in-game lighting condition. However, here we are concerned with the satisfaction, enjoyment and perceived realism of the dynamic effects of blur, latency and lighting for mobile AR games.

To measure the total system latency Jacobs et al. [Jacobs et al. 1997] use a blinking LED and its image displayed on the screen to measure their temporal offset with photo transistors and an oscilloscope. Friston and Steed [Friston and Steed 2014] present and compare a number of latency measurements for VR environments. The delay from a known motion to the display reaction includes tracking, computation and display latencies. In our test cases, it is not necessary to measure motion. Instead, we adapt the two approaches and use a camera to record a modulated light source directly and on the screen at the same time.

Especially for head mounted displays (HMDs) the effect of latency is critical for the user's comfort and Carmack [Carmack 2013] describes a number of strategies to reduce lag. Sielhorst et al. [Sielhorst et al. 2007] measure latency of an AR see through device, and they assert latency is also a crucial factor for a user's task performance, but do not perform a user study to verify this statement.

3 Experiments

We conducted three experiments to assess the impact of different camera and lighting effects in AR games on task performance and perception. Our subjects were 24 to 55 years old and around 80% of them are male.

In the first experiment the users played an AR game *ARTravelers* multiple times with different amounts of artificial camera blur added to the background video image and to the foreground virtual content. The second experiment used the same game with different amounts of added artificial image latency. After each short game, the users answered questions about enjoyment, satisfaction on their resulting score, and perceived realism and immersion.

The AR game used in the first two experiments uses a cubical marker, which is mounted about chest high on a tripod in the center of a room, as depicted in Fig. 1a. For both experiments, each player plays five rounds. In each round, targets spawn randomly around the cube. The player needs to align the mobile device with the target and the cube in order to destroy the target. The cube has a side length of 20 cm and it was mounted on a tripod with about 15 degrees incline. With the cube inclined, the players need to raise and lower the device to achieve high scores, which increases the difficulty of the game. 100 points are awarded for destroying a target with alignment errors up to 1 degree and 30 points for alignment errors of 60 degrees, linearly interpolated in between. Faster shooting allows players to destroy more targets and thus collect more points. The duration of each round was 60 seconds in the first experiment and 40 seconds in the second experiment.

In the third experiment we employed *ARPix*, an AR application that lets someone take a photo of the user together with a virtual character, as depicted in Fig. 1b. Four different versions of the image with different rendering techniques were presented to the user. He or she was then asked to choose the image that looks the most realistic.

In order to statistically evaluate the experiments, we designed them as follows: In the first experiment (blur) the independent variable is the amount of blur, which is nominally defined within three configuration scenarios A, B, and C. The dependent variables where the achieved score (ratio) as well as the answered questions about enjoyment (interval), satisfaction (interval), realism (interval), and matching (interval). The individual questions are discussed in Section 4.1. For the second experiment (latency) the independent variable was the amount of artificially added latency (ratio), whereas the dependent variables were the achieved score (ratio) and the answered questions about enjoyment (interval), satisfaction (interval), and responsiveness (interval).

The games ARTravelers and ARPix were developed in Unity3D[1] using Qualcomm's Vuforia[2] package for AR. They employ natural image marker tracking as performed by Vuforia for the camera pose estimation. ARTravelers ran on an iPad Air (MD785GP/A), iOS version 7.0.3, for the first two experiments, while ARPix ran on an iPad 3rd Generation (MD371FD), iOS version 5.1.1.

[1]http://unity3d.com/, visited 2014-07-14

[2]http://www.qualcomm.com/solutions/augmented-reality, visited 2014-07-14

4 Camera Motion Blur Experiment

Cameras capture an image by exposing the photographic film or digital sensor to photons. Depending on the brightness of the scene being photographed, that is, the number of photons arriving at the film or sensor, the camera settings must be adapted in order to capture a well exposed image. The camera settings comprise the exposure time (shutter speed for video), the film sensitivity or sensor gain, and the aperture of the lens. For darker scenes, as the sensitivity and the aperture are typically more limited (especially in small form cameras as integrated into mobile devices), the exposure time is often increased. This inevitably leads to blurred images if objects in the scene move relative to the camera.

AR applications rely on immersion created through seamlessly mixing the rendered virtual content (usually in the foreground (FG)) with the camera image (usually in the background (BG)). If the FG image does not visually match the BG image, we expected that the immersion may break. Our goal was to test situations where the FG blur matched the BG blur and situations where it did not match. As it is more feasible to add motion blur to the virtual rendered content than to remove it from the camera image, we chose the former method.

In this experiment we chose a bright room, lit by daylight through windows and by lamps (see Fig. 1a). In that way the exposure time is short, resulting in images with very little motion blur. Then, we artificially blurred the BG and the FG independently to simulate a longer exposure time and examined the impact on the task performance and the players' perceptions. The amount and direction of the motion blur depends on the camera motion and is calculated as a translation relative to the marker.

4.1 Experiment Setup

During the experiment, each participant played 5 rounds, 60 seconds each. In each round, a blur configuration scenario was randomly selected. We presented the game to the player without modification (scenario A), with artificial motion blur only added to the BG (scenario B), and with artificially added blur to both the FG as well as the BG (scenario C), see Table 1. Scenario A is the default case in a bright environment. Scenarios B and C simulate a darker scene when the device's camera switches to a longer exposure time. However, in scenario C there is also artificial FG blur added to match the BG blur, which we expected to create a more realistic mixing of virtual and real images than in scenario B. The player was not informed about the intention of the experiment and thus was not aware that there would be artificial blur added to the game.

Table 1: *Configurations scenarios for the motion blur experiment.*

	Foreground Blur	Background Blur
Scenario A	Off	Off
Scenario B	Off	On
Scenario C	On	On

Fig. 2 depicts screenshots from the ARTravelers game, modified for this experiment. During the game, the user sees the cube with targets spawning from its center. These targets have to be destroyed by walking to the target and aligning the device with the target and the marker cube. The head-up display shows the collected points, the remaining time, and the time during which the tracking was lost.

We chose to add a relatively strong blur, compared to the naturally occurring blur in dark settings, because we wanted to investigate if the participants would notice the blur at all and if the blur had any influence on task performance.

After each round, the participants were asked to answer the following questions:

- **Enjoyment:**'How much did you enjoy this run?'
- **Score satisfaction:** 'How satisfied are you with your score in this run?'
- **Realism:** 'How realistic did the game look in this run?'
- **Matching:** 'How well did the virtual content (foreground) match with the camera image (background)?'

Each question could be answered on a scale from 1 (not at all) to 5 (very much).

4.2 Experiment Results

The first out of five rounds was considered a training round and was omitted in the further analysis to remove a strong influence of the learning effect. This learning effect can clearly be observed in Fig. 3b. We recorded 12 participants and thus, 48 rounds in total.

The box plots of all 48 rounds' scores ordered into the blur configuration scenarios in Fig. 3a visually show a slight negative trend from scenario A to B to C. However, the ANOVA yielded no statistically significant connection, as described in Table 2. Thus, the hypothesis that the blur configuration scenario does not influence the player's performance or experience cannot be discarded ($p > 0.05\%$). This may be due to the fact that there were not enough participants recorded or that all participants experienced the blur scenarios very differently. Some players might have looked over the mobile device to orient themselves in the real world and ran directly to the target's position. Others might have been more confused by the blur and scored less.

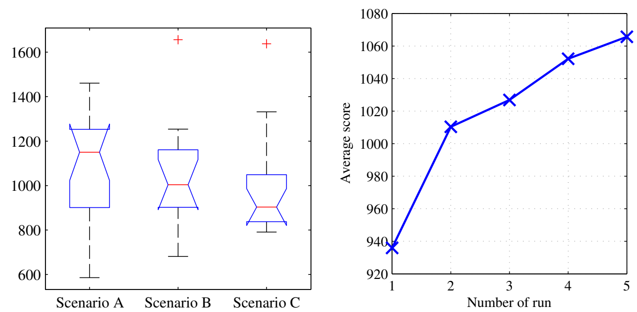
Table 2: *ANOVA statistic for the influence of blur scenarios on player performance (scores) and survey answers (enjoyment, score satisfaction, realism, matching). In our experiment we could not observe a significant influence of the blur scenario on the dependent variables.*

Dependent variable	SS	*df*	MS	F	p
Scores	75761	2	37880	0.64	0.532
Enjoyment	2.115	2	1.057	0.95	0.392
Score satisfaction	2.447	2	1.223	0.78	0.463
Realism	0.731	2	0.365	0.48	0.620
Matching	1.066	2	0.533	0.59	0.560

Fig. 4 depicts the correlation matrix for the the independent variable, the blur configuration scenario, and the dependent variables, the scores and the answers. Confirming the results from Table 2, none of the independent variables correlates in a statistically significant way with the dependent variable. Interestingly, the correlation between the scores and the enjoyment is also not a significant correlation, even though intuitively, one would expect a correlation. However, there seems to be, next to the obvious correlation between enjoyment and scores satisfaction, a strong ($r > 0.4$) correlation between enjoyment and matching as well as very strong ($r > 0.69$) correlation between enjoyment and realism. We observe that, independent of the achieved scores and how the blur was actually configured, the participants answered that they enjoyed the last round if they voted for a high perceived realism of the game as well as a high matching of BG and FG.

(a) *A: FG blur = Off; BG blur = Off* **(b)** *B: FG blur = Off; BG blur = On* **(c)** *C: FG blur = On; BG blur = On*

Figure 2: *ARTravelers in-game screenshots with different blur configuration scenarios (A, B, C) during the Camera Motion Blur experiment.*



(a) **(b)**

Figure 3: *(a) depicts box plots of scores for each scenario. Visually, a decreasing trend is noticeable from scenario A to B to C. That is, the stronger the total blur, the less the players scored. However, the trend is not statistically significant. (b) shows the average score per round for all players. A learning effect is clearly visible.*

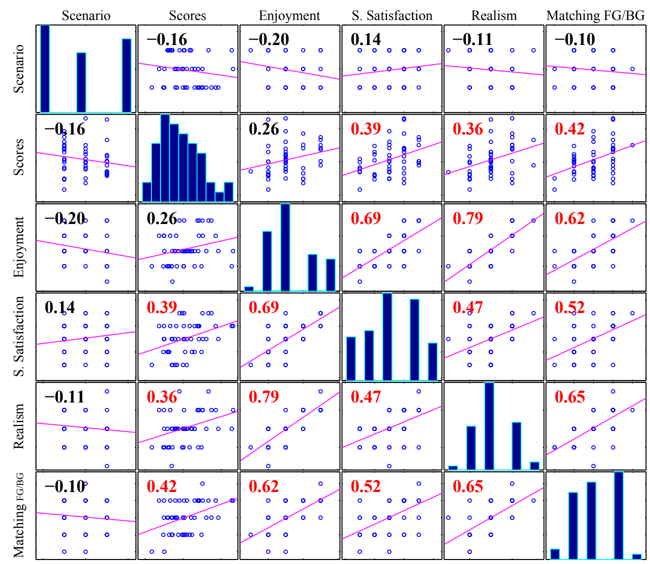


Figure 4: *Correlation matrix for the independent variable (scenarios) and all dependent variables (scores and answers). A red correlation value indicates a statistically significant correlation.* Scena *refers to the scenarios (A, B, C, from left to right and bottom to top),* scores, enjoyment, satisfaction, realism, *and* matching *(1 to 5, from left to right and bottom to top) refers to the recorded performance and answers of the participants. The plots on the diagonal depict a histogram for each distribution.*

5 Latency Experiment

In video see-through AR games and applications running on mobile devices and HMDs the virtual content is rendered onto the camera image. Until the final image is visible on the device's display, processing, synchronization and signal transmissions cause delay and add up to the total system latency.

This latency is an important factor for reactive games as well as for augmented reality applications. We used the same game as in the motion blur experiment (Section 4) to investigate the influence of latency on the task performance.

In this experiment, we artificially increase the latency of the frames presented to the players and record again task performance as well as usage experience. We hope to gain insight about the degree to which latency influences these measures in order to judge the importance of reducing sources of latency.

5.1 Native System Latency Measurements

To measure the total native latency caused by the hardware (camera) and software (operating system and middleware) in ARTravelers, we place blinking LEDs in the view of the device's camera and record both the LEDs on the screen as well as the LEDs directly with a 200 Hz camera, as depicted in Fig. 5. This approach is similar to Jacobs et al.'s [Jacobs et al. 1997] method.

We then analyze the brightness of both LEDs' image regions in the recorded video over time by cross correlating the two pixel brightness signals (thresholded) to find the system latency. In Fig. 6, the two pixel brightness signals for an example measurement are shown. To get a stable signal, an average of a 10×10 pixel area is used. To avoid a measurement error from rolling shutter, which is used in both cameras, we tried to vertically center the lamps in the image. A textured background and the marker cube were also visible to the device's camera in order to measure a realistic latency like during the game. The offset between the two signals is clearly visible. In this measurement it is about 100 ms.

We conducted measurements for several settings that each had different influences on the latency, summarized in Table 3. The lighting column denotes how bright the scene was lit. It is important

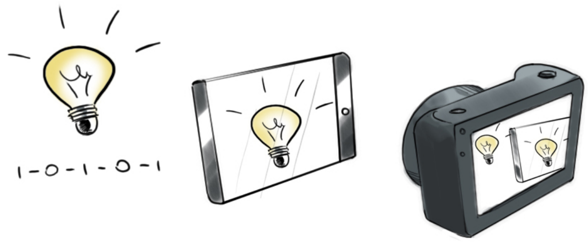
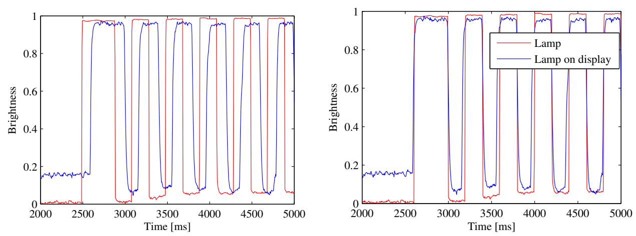


Figure 5: *Latency measuring setup. The mobile device running ARTravelers is pointed at blinking LEDs. A high-speed camera captures both, the LEDs and the LEDs on the display. The recorded video can then be analyzed to calculate the delay between the LEDs and the LEDs on the display.*



(a) **(b)**

Figure 6: *Two brightness signals over time. In (a) the offset between the signals is clearly visible, about 100 ms. In (b) the offset has been calculated and removed from the signals with the method discussed in Section 5.1.*

because low light conditions may lead to increased exposure time and therefore introduce larger latencies. The features column denotes if the device's camera was pointed at a background with a high number of image features (e.g. SIFT features [Lowe 2004]). With more potential features, the feature matching takes longer as there are more features with which to compare. The application column refers to the application that was running while recording the measurement video.

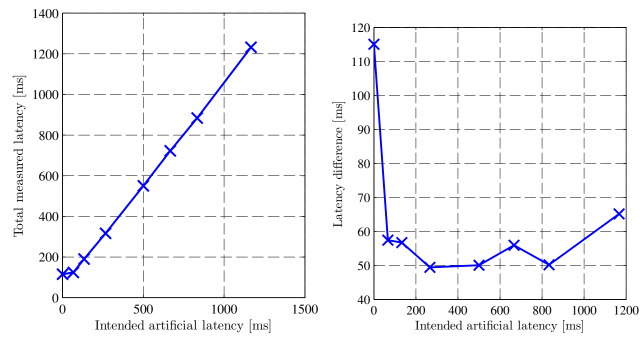
Measurement A used the camera app preview and bright lighting to measure the minimal possible latency (around 90 ms). The conditions in B, C and D were chosen to cover different situations or setups for our game. In normal lighting (B, C) the latency is around 100 ms and in a low lighting situation it grows to around 130 ms.

Table 3: *Measured native system latency for different settings.*

Setting	Lighting	Features	App	Avg latency
A	bright	-	iOS camera app	89.9 ms
B	normal	high	ARTravelers	101.5 ms
C	normal	low	ARTravelers	97.3 ms
D	dark	low	ARTravelers	131.1 ms

5.2 Artificial Latency Measurements

We modified ARTravelers such that it delays all shown frames by N frames to introduce additional artificial latency (AAL), that is, latency additional to the existing native system latency. The application runs at constant 30 fps and therefore every AAL frame adds an additional 33.3 ms latency to the native latency. To verify, we measured the AAL using the same method as in Section 5.1. In Fig. 7a the measured total latency over the intended artificial latency is plotted. As expected from the results in Table 3, there is a minimal total latency of about 115 ms. For higher AAL, the native latency decreases to about 50 ms to 60 ms.



(a) **(b)**

Figure 7: *(a) depicts the total measured latency over the intended AAL. (b) shows the native latency (i.e. the difference between the total measured latency and the intended AAL) over the intended AAL.*

Fig. 7b depicts the native latency (i.e. the difference between the total measured latency and the intended AAL) over the intended AAL. In the experiment, this curve is used as a look up table to calculate the true total latency based on the intended AAL.

5.3 Experiment Setup

Similar to the blur experiment, each participant played multiple rounds of ARTravelers with different amounts of AAL and tried to achieve as high a score as possible. The participant was not informed about the intention of the experiment, i.e. that there will be increased latency in the game. First, the participant played two introductory rounds: one without AAL and one with a high (333 ms) AAL. These two introductory rounds were not used for further analysis. Then the participant played another five rounds, each lasting 40 seconds, with AAL bucket randomly assigned to each round. We intended to sample the lower AAL space denser and defined the buckets as follows:

- **Bucket 1:** 0 to 1 frames (0 ms to 33 ms) AAL
- **Bucket 2:** 2 to 4 frames (66 ms to 132 ms) AAL
- **Bucket 3:** 5 to 9 frames (165 ms to 297 ms) AAL
- **Bucket 4:** 10 to 16 frames (333 ms to 528 ms) AAL
- **Bucket 5:** 17 to 35 frames (561 ms to 1155 ms) AAL

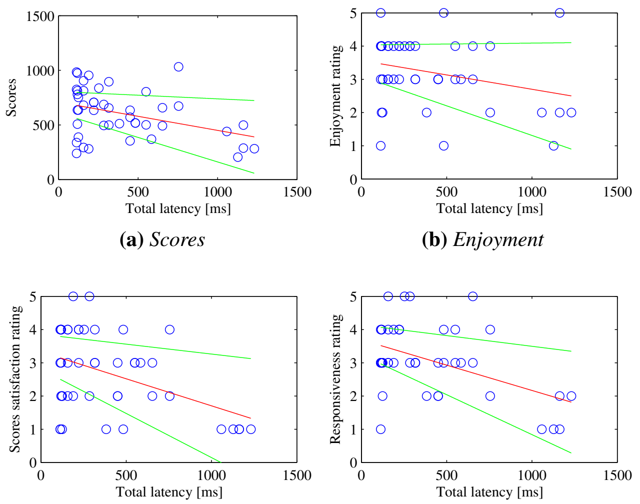
Over the five rounds, each player was assigned each bucket once. Inside each bucket a random value was generated for each player. After each round, the participants were asked to answer the following questions:

- **Enjoyment:** 'How much did you enjoy this run?'
- **Score satisfaction:** 'How satisfied are you with your score in this run?'
- **Responsiveness:** 'How responsive did the game feel?'

Each question could be answered on a scale from 1 (not at all) to 5 (very much).

5.4 Experiment Results

We used a linear regression model to explain the connection between latency, as the independent variable, and enjoyment, score satisfaction, and responsiveness, as dependent variables.



(a) *Scores* **(b)** *Enjoyment*

(c) *Score satisfaction* **(d)** *Responsiveness*

Figure 8: *Linear regressions of scores, enjoyment, scores satisfaction and responsiveness over total latency. The red lines depict the fitted linear curves and the green lines define the lower and upper limit of the 95% confidence interval.*

Table 4: *Regression analysis results for the dependent variables scores, enjoyment, score satisfaction, and responsiveness over the independent variable, the total latency.*

Dependent variables	R^2	F	p
Scores	0.147	7.40	0.009
Enjoyment	0.077	3.59	0.064
Score satisfaction	0.191	10.13	0.002
Responsiveness	0.221	12.21	0.001

The experiment showed a significant ($p < 0.01$) connection between latency and scores as well as between latency and responsiveness. As expected, but as opposed to the blur experiment, the player could perceive a delay in the game and at the same time his or her performance suffered during high latency rounds, see Fig. 9 and Table 4. However, latency had no significant influence ($p > 0.05$) on the reported enjoyment of the participants. Even in ARTravelers, which requires high concentration but little reaction skills, a negative effect of latency on task performance is apparent.

6 Realistic Lighting and Camera Artifacts Experiment

As discussed in [Wood et al. 2004], the success of video games depends on a high degree of visual realism. High-quality realistic graphics were rated as important by four-fifths of the participants in their comprehensive study.

In this experiment we investigated the impact of different mixing techniques on the users' perception of realism for AR applications. ARPix is an application that lets a person take a picture of another

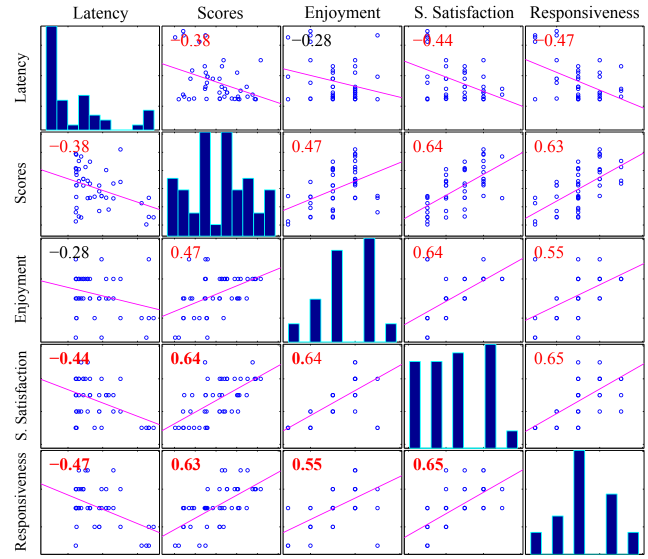


Figure 9: *Correlation matrix for the independent variable (latency) and all dependent variables (scores and answers). A red correlation value indicates a significant correlation.* Laten *refers to the latency (0 ms to 1300 ms, from left to right and bottom to top),* scores, enjoyment, satisfaction, *and* responsiveness *(1 to 5, from left to right and bottom to top) refers to the recorded performance and answers of the participants. The plots on the diagonal depict a histogram for each distribution.*

person posing with an augmented virtual character, Eva, as depicted in Fig. 1b. The goal of ARPix is to blend Eva into the camera image with a high degree of realism. To achieve this goal, two effects are integrated:

A diffuse sphere and a specular sphere mounted at each side of the image marker reflect the lighting condition of the current real scene. ARPix, knowing the positions of the spheres, calculates the lighting condition and applies it to the virtual scene.

Cameras in cell phones and tablets often have a limited lens and sensor quality, resulting in slightly blurry, distorted, and noisy images. These effects can be summarized as camera artifacts. To enhance the realism of our virtually rendered images, artificial camera artifacts (ACA) are added. In our experiment, with enough light, sensor noise was not an issue, and distortion was not visible as Eva is positioned close to the center of the frame. The ACA in the experiment only included artificial blur. The radius of the Gaussian blur filter was manually adjusted to best match the camera image.

6.1 Experiment Setup

We installed ARPix, including the marker and the two spheres, and positioned a bright white/blueish lamp to the left side and a smaller reddish lamp to the right side of the user, as depicted in Fig. 1b. We conducted this experiment at two different locations (1 and 2), with different participants and similar lighting conditions. The participants at location 1 were less experienced in visual computing and games as the participants at location 2.

After taking the picture with the participant, four different image versions with different combinations of the aforementioned effects were generated and presented to the participant. Table 5 describes the mapping from the scenario to the contained effects. Correct lighting refers to the lighting captured from the spheres. For the incorrect lighting, we simply created a light shining from the bottom up at Eva.

Table 5: *The effect configurations for the four images presented to the user.*

Scenario	Lighting	Artificial Camera Artifacts
A	Correct	On
B	Correct	Off
C	Incorrect	On
D	Incorrect	Off

For comparison, all four images were presented to the user with a randomized position on the screen. The user was then asked to choose the most realistic image (1 to 4). Fig. 10 depicts a screenshot taken from the image selection screen during the running application. The red letter was added after the screen was captured to indicate the scenario.

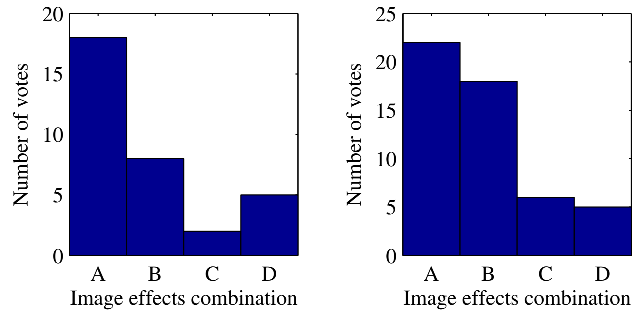
Figure 10: *Screenshot from ARPix. Here, the user can select his or her preferred version of the image taken with the virtual character. In 1 and 4 the virtual character is correctly lit and in 1 and 2 she is slightly blurred.*

6.2 Experiment Results

The preferred image votes for both locations are depicted in Fig. 11. In both locations the participants mainly voted for the image with correct lighting and enabled ACA (scenario A), which confirms our expectation. In location 2 a greater part of the participants voted for configuration B compared to location 1. We believe that this is due to the fact that the percentage of technical people that are working in the field of visual computing was significantly higher in location 2 than in location 1. Participants with a visual computing background may be more sensitive to image effects and may recognize the slightly blurry virtual content and consider it a visual defect, whereas less technical participants would not directly spot the blur but unconsciously feel that the image with ACA enabled blends better into the background.

7 Conclusion

Consistent with an earlier study of motion blur on player's experience of a racing game [Sharan et al. 2013], we observed that even strong motion blur in our mobile Augmented Reality game,



(a) *Location 1* **(b)** *Location 2*

Figure 11: *Votes for the most realistic picture from a selection of four effect combination scenarios A: correct lighting and artificial camera artifacts, B: correct lighting and no artificial camera artifacts, C: incorrect lighting and artificial camera artifacts, D: incorrect lighting and no artificial camera artifacts.*

ARTravelers, may be noticed by players but does not affect the player's performance or feeling of immersion.

This was also true in our variations of applied foreground and background blur, specific to AR scenarios. In common with many console action games, our AR game is demanding and requires continuous player attention. Most players, when asked if they noticed differences from round to round, would answer that they were too distracted focusing on the game and did not pay any attention to effects. For such intense games, the players are less likely to notice such camera image composition effects.

Given that immediate camera image synthesis reality mixing effects, such as motion blur, did not impact enjoyment, we further assessed an alternative animation related factor, latency. As the role of latency in mobile AR applications had not previously been dealt with in terms of overall impact in the user experience, we introduced a latency measuring method based on [Jacobs et al. 1997] and applied a range of additional exaggerated artificial frame delays. This resulted in a strong correlation between lower latency and positive user experience. Significant observable impact of reduced latency for AR mobile games was recorded on realism, enjoyment, satisfaction, matching and score. The strongest of these was realism, and overall, we validate that low-latency is critical to the sense of presence and engagement in mobile AR games.

Finally, we investigated the impact of visual realism in mobile AR for the category of applications where aesthetic quality is a factor. In this test, we measure the impact of responding to dynamic changes in the lighting environment and camera sensor resolution blur matching. Here we found a strong preference to the inclusion of more realistic lighting environment matching, but no noticeable preference for resolution matching between the camera and the rendered augmented reality graphics.

In summary, this article has presented a series of user experience evaluations of dynamic augmented reality blending approaches. For image effects we detected no significant impact on the user experience for mobile AR games. However, latency was highly significant in this domain. Where the entertainment application domain deals with the importance of aesthetic quality, such as in a self portrait AR photo app, the realism oriented reality mixing measures were found to be important.

In our targeted set of experiments, we have shown several interesting observations and conclusions. Latency clearly warrants further scruntization, as AR systems improve performance, the effects of varying latencies below 50 ms is particularly interesting as signalled by this study. More generally, this work motivates further research, such as a more extensive study of the influence of motion characteristics across different genres of AR games.

Acknowledgements

This work was funded by the EU within the FI-CONTENT 2 initiative. We would like to thank our artists, Alessia and Mauri, for the great help with drawings and videos, our participants for their effort playing the games, and foremost our reviewers for their time and helpful comments.

References

AGUSANTO, K., LI, L., CHUANGUI, Z., AND SING, N. W. 2003. Photorealistic rendering for augmented reality using environment illumination. In *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, Washington, DC, USA, ISMAR '03, 208–.

AZUMA, R. T. 1997. A survey of augmented reality. *Presence 6*, 4, 355–385.

BAE, H., FOWLKES, C. C., AND CHOU, P. H. 2013. Accurate motion deblurring using camera motion tracking and scene depth. In *Proceedings of the 2013 IEEE Workshop on Applications of Computer Vision (WACV)*, IEEE Computer Society, Washington, DC, USA, WACV '13, 148–153.

CARMACK, J., 2013. Latency mitigation strategies. `http://www.altdev.co/2013/02/22/latency-mitigation-strategies/`.

FISCHER, J., BARTZ, D., AND STRASSER, W. 2006. Enhanced visual realism by incorporating camera image effects. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, Washington, DC, USA, ISMAR '06, 205–208.

FRISTON, S., AND STEED, A. 2014. Measuring latency in virtual environments. *IEEE Transactions on Visualization and Computer Graphics 20*, 4 (Apr.), 616–625.

JACOBS, M. C., LIVINGSTON, M. A., AND STATE, A. 1997. Managing latency in complex augmented reality systems. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics*, ACM, New York, NY, USA, I3D '97, 49–ff.

JOSHI, N., KANG, S. B., ZITNICK, C. L., AND SZELISKI, R. 2010. Image deblurring using inertial measurement sensors. In *ACM SIGGRAPH 2010 Papers*, ACM, New York, NY, USA, SIGGRAPH '10, 30:1–30:9.

KANBARA, M., AND YOKOYA, N. 2002. Geometric and photometric registration for real-time augmented reality. In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, Washington, DC, USA, ISMAR '02, 279–.

KLEIN, G., AND MURRAY, D. W. 2010. Simulating low-cost cameras for augmented reality compositing. *IEEE Transactions on Visualization and Computer Graphics 16*, 3, 369–380.

KNEZ, I., AND NIEDENTHAL, S. 2008. Lighting in digital game world : Effects on affect and play performance. *Cyberpsychology & Behavior 11*, 2, 129–137.

KRUIJFF, E., SWAN, J., AND FEINER, S. 2010. Perceptual issues in augmented reality revisited. In *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on*, 3–12.

LOWE, D. G. 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision 60*, 2 (Nov.), 91–110.

OH, S., AND KIM, G. 2014. Robust estimation of motion blur kernel using a piecewise-linear model. *Image Processing, IEEE Transactions on 23*, 3 (March), 1394–1407.

OLIVEIRA, J., FIGUEIREDO, M., AND BIOUCAS-DIAS, J. 2014. Parametric blur estimation for blind restoration of natural images: Linear motion and out-of-focus. *Image Processing, IEEE Transactions on 23*, 1 (Jan), 466–477.

PARK, Y., LEPETIT, V., AND WOO, W. 2012. Handling motion-blur in 3d tracking and rendering for augmented reality. *IEEE Transactions on Visualization and Computer Graphics 18*, 9 (Sept.), 1449–1459.

SHARAN, L., NEO, Z. H., MITCHELL, K., AND HODGINS, J. K. 2013. Simulated motion blur does not improve player experience in racing game. In *Proceedings of Motion on Games*, ACM, New York, NY, USA, MIG '13, 149:149–149:154.

SIELHORST, T., SA, W., KHAMENE, A., SAUER, F., AND NAVAB, N. 2007. Measurement of absolute latency for video see through augmented reality. In *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*, 215–220.

SWAN(II), J. E., AND GABBARD, J. L. 2005. Survey of user-based experimentation in augmented reality. In *Proceedings of 1st International Conference on Virtual Reality, HCI International 2005*.

TAO, M. W., MALIK, J., AND RAMAMOORTHI, R. 2013. Sharpening out of focus images using high-frequency transfer. In *Computer Graphics Forum*, vol. 32, The Eurographics Association and Blackwell Publishing Ltd., 489–498.

WOOD, R. T. A., GRIFFITHS, M. D., CHAPPELL, D., AND DAVIES, M. N. O. 2004. The structural characteristics of video games: A psycho-structural analysis. *Cyberpsy., Behavior, and Soc. Networking 7*, 1, 1–10.