A Framework for Perceptual Studies in Photorealistic Augmented Reality

Martin Knecht
Institute of Computer Graphics and Algorithms
Vienna University of Technology

Andreas Dünser
HIT Lab NZ
University of Canterbury

Christoph Traxler
Institute of Computer Graphics and Algorithms
Vienna University of Technology

Michael Wimmer
Institute of Computer Graphics and Algorithms
Vienna University of Technology

Raphael Grasset
HIT Lab NZ/ICG
University of Canterbury
Graz University of Technology

Abstract

In photorealistic augmented reality virtual objects are integrated in the real world in a seamless visual manner. To obtain a perfect visual augmentation these objects must be rendered indistinguishable from real objects and should be perceived as so. We propose in this paper a research testbed framework to study the different unresolved perceptual issues in photorealistic augmented reality and its application to different disciplines. The framework is able to compute a global illumination approximation in real-time and therefore leverage a new class of experimental research topics.

Keywords: Human perception, photorealistic augmented reality, real-time global illumination

Index Terms: H.1.2 [Models and Principles]: User/Machine Systems—Human factors; H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Evaluation/methodology

1 Introduction

Augmented Reality (AR) technology offers a way to represent visually virtual content related to the real world. Its application have been proposed to advertise products, in architectural visualization, edutainment systems or for enhancing cultural heritage sites.

As lot of progress have been considering the spatial registration of real and virtual content (geometric), the visual integration (photometric) is still confronted with a large number of issues. These problems can be divided into two main areas: the ones which are of technical nature, like the narrow field of view of Head Mounted Displays (HMDs) and the other ones which are of perceptual nature. For example depth perception differs for virtual objects compared to real ones. Although there is a large number of studies in this area, there are still open questions and we are not absolutely certain which parameters influence perception.

To address these issues, we are introducing in this paper a software research framework offering new possibilities to investigate these perceptual issues. With the proposed framework, we are able to study perceptual issues with shadows, dynamic environmental illumination and indirect illumination as shown in Figure 1 – all at real-time frame rates.. Kruijff et all. [1] wrote an a taxonomy of the main perceptual issues in AR. They classified them grounded on the so called perceptual pipeline which consists of five stages: *Environment*, *Capturing*, *Augmentation*, *Display Device* and finally the *User*. The work in progress, we present here, fits into the *capturing* and *augmentation* stages of the perceptual pipeline. Our main contributions are:

- A framework to study photorealistic rendering techniques in AR to investigate perceptual issues and visual cues
- Advanced rendering system that enables different rendering modes and styles
- A preliminary user-study to test our framework

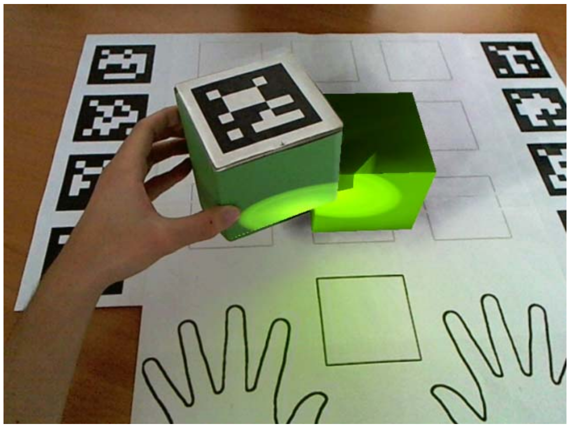

Figure 1. This figure shows the augmented scene of our experiment including shadows and color bleeding.

2 Related Work

We divided the related work section into three main parts. In the first, we show a selection of work that was done in the area of perception regarding to shadows and indirect illumination. Second we show two studies about the perception of environmental illumination and third we show work that is directly related to our preliminary user-study and the proposed framework.

A lot of research studies the influence of shadows and indirect illumination in augmented - and virtual reality applications. Hubona et al. [2] experimented with positioning and resizing tasks under varying conditions. They found significant differences for all independent variables. Sugano et al. [3] studied how shadows influence the presence of virtual objects in an augmented scene. The experiments showed that the shadows increased the presence of the virtual objects. Madison et al. [4] generated several different images of a plane and a cube. With different visual cues enabled and disabled the participants had to tell whether the cube was touching the plane or not. Similar to that work, Hu et al. [5] generated several different images of a plane and a large box using a Monte-Carlo path tracer. Their results showed that stereo vision is a very strong cue followed by shadows and indirect illumination. Furthermore shadows combined with indirect illumination are similarly as strong as stereo vision.

In all of these studies, indirect illumination was either not included as an independent variable or the studies used static images to overcome the computational costs caused by indirect illumination. However, the proposed research framework makes it possible to setup interactive experiments including indirect illumination effects.

Some studies investigate thresholds in environmental illumination. Nakano et al. [6] studied how much the resolution of an environment map could be decreased until the increasing error is noticeable. Lopez-Moreno et al. [7] studied how much the illumination direction of an object could differ until human observers noticed the error. The results showed that the error threshold was even larger in real scenes than in synthetic ones. However, only static environments were used for these experiments and it would be interesting how the thresholds behave in dynamic setups.

Our research framework is an extension of the method proposed by Knecht et al. [8]. It basically uses a variation of the instant radiosity algorithm by Keller [9] combined with differential rendering from Debevec [10] to compute global illumination suitable for augmented reality applications.

Similar to our study Thompson et al. [11] tried to find out if improved rendering methods also improve distance judgment. The setup of the experiment and distances to estimate are different to our preliminary user-study. However, the results look similar to ours (see Section 6.3).

3 PHOTOREALISM IN MIXED REALITY

As written in Section 1 it sounds plausible that virtual objects should look photorealistic in an augmented reality setup. In the ideal case virtual objects are indistinguishable from real ones. However, what does it take to make virtual objects look photorealistic and even better, make them indistinguishable from real objects? We start with the work from Ferwerda [12]. He introduced three different varieties of realism and pointed out, that an image is just a representation of a scene. This representation describes selected properties and we should not confuse this with the real scene. The three varieties are:

Physical realism, where the visual stimulus of a scene is the same as the scene itself would provide. Physical realism is hard to achieve due to the lack of appropriate display devices that can recreate the exact frequency spectrum.

Photo-realism, where the visual response is the same as invoked by a photograph of the scene. This kind of realism should be targeted in photorealistic augmented reality systems based on video-see-through output devices. If it would be possible, that the virtual objects are represented using the same kind of photorealistic mapping function, they would be indistinguishable from real objects.

Functional realism, provides the same visual information as the real scene does. That means, that the image itself can be rather abstract but the information retrieved from it, is the same. A construction manual of a cupboard will contain abstract drawings but normally no photographs for example.

3.1 Studies on photorealism

Having Ferwerda's [12] three varieties of realism helps to focus on what kind of realism we want to achieve in photorealistic augmented reality. However, it is still not fully understood what photo-realism actually means in a perceptual context. Therefore Hattenberger et al. [13] conducted experiments to find out, which rendering algorithm creates the most photorealistic images. They used a real scene and added a virtual cow in the middle of it. Several different rendering algorithms were used to calculate the final results. Observers had to choose between two images compared to a photograph of the scene and decide which one looks more real. Results showed that observers preferred light simulations that took indirect illumination into account and furthermore, that noisier images were preferred to more smooth ones (with some exceptions). Although the authors state, that the results cannot be generalized because they belong to this particular scene, the results indicate, that there are also other important factors in photorealistic augmented reality that influence the perception of the scene.

Elhelw et al. [14] tried a different approach. They used an eye-tracking system to find the gaze points in images. From that they derived which image features were important for the participants to decide if the image looks real or not. They found light reflections/specular highlights, 3D surface details and depth visibilities to be very important image features. For the user-study they used different sets of images from clinical bronchoscopy. These images look quite abstract in shape and texture. However, it would be very interesting to test this method on other images that are related to augmented reality applications.

These are two examples of user-studies that tried to find answers on what makes an image photorealistic, without altering specific image features. We propose to divide the known image features in an augmented reality setup into two main categories: The *visual cues* described in Section 3.2 and the *augmentation style* described in Section 3.3. While visual cues have a local nature augmentation style can be seen as global features in an image.

3.2 Visual Cues

Visual cues are very important for the human visual system (HVS) as they help to organize and perceive the surrounding environment. Visual cues can deliver depth information and let us recognize inter-object relationships.

In augmented reality visual cues can be exploited to embed virtual objects into the real scene. We split visual cues into *inter-object spatial cues* and *depth cues*.

Inter-object spatial cues

Shadows belong to the strongest spatial cues available. They define a spatial relationship between the shadow caster and the shadow receiver. The influence of shadows was studied in several experiments (see Section 2). Rademacher et al. [15] furthermore found out, that the characteristics of soft-shadows changed the perceived realism in images.

Like shadows *indirect illumination* between objects defines a spatial relationship. Although inter-reflections are not a strong cue as shadows are, their influence is still significant [4].

Depth cues

Beside spatial cues like shadows or indirect illumination, cues that serve as a source for depth information are of particular interest, as these allow reconstructing our surrounding environment. Drascic and Milgram [16] as well as Cutting [17] presented a list of depth cues. The cues can be divided into four main groups: Pictorial depth cues, kinetic depth cues, physiological depth cues and binocular disparity cues.

Pictorial depth cues are features, which give information about the objects position in a still image. Such cues can be occlusion, linear perspective, relative size, texture perspective or aerial atmospheric perspective.

Kinetic depth cues provide information through change of the viewpoint or moving objects. Relative motion parallax and motion perspective (falling raindrops – near vs. far) are two examples. Another one is the so called kinetic depth effect. Imagine a point cloud that rotates around its upper axis. The structure of the point cloud is easily recognized. However if the cloud stops rotating every point falls back into the screen plane and the structure is not visible anymore.

Physiological depth cues deliver information to the HVS about the convergence and accommodation of the eyes.

Binocular Disparity is another depth cue that is similar to the motion parallax depth cue. The HVS automatically transforms the disparity seen due to our two eyes into depth perception. Obviously this cue only exists when a stereo rendering setup is used in experiments.

3.3 Augmentation Style

Beside visual cues that should be supplied by the rendering system it is also important that the augmentation style of virtual objects is similar to the visual response of the scene. Kruijff [1] mentioned several areas where perceptual issues may arise.

Illumination

Virtual objects that are rendered into the captured image of the real world must be illuminated correctly. This is often done by using a chrome sphere to capture the incident illumination at the point where the objects will be placed. This method belongs to the outside-in approaches. Debevec [10] introduced a way to use several images with different exposure times to create a high dynamic range (HDR) environment map. However, this process is time consuming and only leads to a static environment map. Inside-out methods instead use a camera with a fish-eye lens to capture the surrounding hemisphere. These methods allow for dynamic environments. Unfortunately there are only a few HDR cameras on the market. So the source for the incident illumination is only of low dynamic range. Once the environment map is acquired, image based lighting methods can be used to illuminate the virtual objects.

Color and Contrast

Currently most cameras offer only a limited color gamut and contrast. These limitations lead to wrong color and contrast representations. A special problem due to this tone-mapping arises, when two different cameras are used; one for video-see through and one to capture the surrounding illumination. Both map the high dynamic range illumination into a low dynamic range, *but* with different tone-mapping functions resulting in wrong colors in the final composed image.

Tone-mapping

The ideal setup for a photorealistic augmented reality system would consist of two equal HDR cameras for video-see-through and environment capturing. Using these two cameras with the same configuration would make the virtual objects look correctly illuminated and there would be fewer errors from the capturing stage. Then the whole rendering process could be performed in HDR and ideally the resulting images would be presented on a HDR display. As we do not have a HDR display our framework uses a tone-mapping operator developed by Reinhard et al. [18] which can be implemented directly on the graphics hardware.

Camera Artifacts

Computer generated images normally look absolutely clean/perfect and do not suffer from artifacts like noise or blurred edges. However, since we embed the virtual objects into a captured video frame, we need to add these artifacts to the virtual objects; otherwise they will be immediately recognized as not being real. Klein and Murray [19] developed a method that imitated a couple of artifacts such as Bayer pattern approximation, motion blur or chromatic aberration. Fischer et al. [20] could improve visual fidelity by removing aliasing artifacts and adding synthetic noise to the rendered objects. These artifacts greatly increase the appearance of the virtual objects.

4 A Research Framework for Photorealistic AR

With the above mentioned information and with the goal to perform experiments in mind, an ideal research framework for photorealistic augmented reality has the following primary requirements:

- It must be very flexible to configure how everything is rendered
- It must be able to produce photorealistic results including augmentation artifacts, so that virtual objects are indistinguishable from real objects.

The framework should make it possible to easily hook in different modules into the rendering pipeline and it should be fast to setup experiments. The API should be designed in a way that new hardware devices can easily incorporate into the existing framework. Furthermore utility functions for data logging, tracking and calibration should be provided.

Such a framework could be used to find out more about how the HVS processes images and how different visual cues alter the perception. Especially in medical AR training simulators it is important that the spatial perception correlates with the real world. Otherwise the students are able to perform the surgery in a simulator, but would have problems in a real world environment.

With these goals in mind we developed a research framework based on the method introduced by Knecht et al. [8]. This method is able to simulate the mutual light interaction between real and virtual objects in real-time. The proposed research framework is developed in C# and runs on Windows 7 64-Bit. The graphical output is done via SlimDX and DirectX 10 APIs. It should be therefore very easy and fast to develop new experiments, as C# offers a lot of tools and functions.

The central object of the framework is a so called *scene* object that is in its main function a hash table to store all the necessary objects for the rendering and serves as a communication platform to pass data from one task to the next. Tasks are pieces in the rendering pipeline that will be executed once every frame. The current framework has several tasks like video capturing, tracking, and rendering. As an example, the video capture task captures a new frame from a camera and passes it to the scene object. When the tracker task is executed it takes the frame, stored in the scene's hash table and uses it for estimating a camera pose. If a new experiment is designed the main procedures of the experiment are methods of an object that implements the specific task interface.

To allow for a very flexible framework the rendering pipeline can be defined in a XML configuration file that can be loaded over the GUI. This way it is easily possible to exchange a tracking system or change a camera without the need to alter the whole experiment.

As a lot of studies are about rendering visual features, shader development should be very efficient. In our framework they can be manipulated in an external editor during run-time. As soon as the shader is saved it will be reloaded automatically. This way instant visual feedback is provided.

The current renderer supports two types of shadows. For spotlight sources we use standard shadow mapping and for indirect illumination we use by default ISMs for every virtual point light. However, standard shadow mapping can also be used for the virtual point lights. Furthermore shadowing and indirect illumination can be switched on and off separately during run-time. In this way the influence of local illumination versus global illumination in an AR setup can be investigated in interactive experiments.

The fish-eye camera currently in use is only able to capture low dynamic range images. However, the rendering framework uses the method from Landis [21] to extrapolate a high dynamic range image from it. This is a very rough approximation and the best solution would be to have a HDR camera.

Dynamic spotlights are also supported. They can either be real pocket lamps that are tracked or simply virtual ones. They will illuminate the real and virtual objects accordingly.

The framework can handle multiple camera streams on the fly and the captured frames are available as textures in the video memory or directly in the main memory. This way they can easily be altered if necessary in a post-capture step.

The tracking interface currently supports three different types of trackers. The first one is the Studierstube Tracking framework. The second one is based on the PTAM tracking method from Klein and Murray [22] and the third one supports the VRPN protocol.

5 Technical Issues

As this is all work in progress there are still several limitations and technical issues that are unsolved. One of the main issues for further perceptual studies is, that the framework in the current stage does not support stereo rendering. This is definitely a goal for future work.

Calibration is crucial when it comes to accurate rendering. As Kruijff [1] mentions there are several points in the perceptual pipeline where errors lower the quality of the final results and this is also true for this framework. If the tracking is not accurate wrong edges are far more visible due to artificial indirect illumination overlays. Methods like the one from Klein and Drummond [23] should be used to accurately move rendered edges to where they are shown in the video stream.

The fish-eye lens camera does not deliver any distance information of the environment. So it is not possible to take near light-sources accurately into account, except they are tracked.

The method used to compose the final images, limits the framework to video see-through HMDs. Furthermore the real-time global illumination computation needs a powerful graphics card and thus mobile augmented reality is not supported yet.

Several different tone-mapping operators exist and each camera has an individual way to map the incident HDR illumination into low dynamic range. This introduces a lot of problems, when compositing the final images and needs manual fine tuning to get satisfying results.

6 Preliminary User-Study

To test our system we have conducted a preliminary user-study on the influence of shadows and indirect illumination for five different tasks.

6.1 Experiment setup

The experiment was conducted at the HIT Lab NZ. The study setup as shown in Figure 2 consisted of a table plate with several BCH markers, two standard USB webcams, a HMD, and two targets (small green cubes with tracking markers). To track hand-movement for task four and five we attached three different markers on the participant hand: One at the index finger, one at the thumb and one at the wrist (see Figure 3).

Figure 2. Participant performing the experiment.

One webcam was attached to the HMD to capture the participants view. The other one was placed above the table. Using this setup we could achieve correct tracking even in situations when the markers were not visible to the head mounted camera.

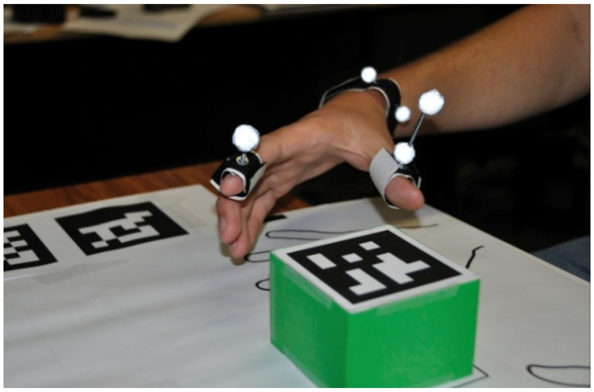

Figure 3. The green box and the markers for tracking the hand position and pose

6.2 Task description

The first task showed a virtual cube at a random position, while the real cube was fixed in the middle of the table. The participants had to estimate the distance between the real and the virtual cube in centimeters.

In the second task the virtual cube was randomly placed in front of the participants. They had to grab the real cube, located on a fixed starting point, and move it to the virtual cube's position. The participants were instructed to perform the remaining tasks (2 - 5) as fast and as accurate as possible.

The third task was similar to task two but this time, the virtual and the real cube were swapped. The real cube was placed at random positions on the table by the experimenter and the virtual cube had to be moved to the same position using a computer keyboard.

In task four the real cube (without any virtual augmentation) was placed at a random position on the table and the participant had to grab and lift it up as fast as possible. Before the task started and the scene was seen through the HMD the participants were asked to place their hands at a fixed starting position.

Task five was similar to task four except that the cube was overlaid with a virtual cube. This way the visual input was virtual, but the tactile input when grabbing and lifting was real.

Rendering modes

For all tasks, we had three conditions (see Figure 4). The first rendered the scene without any cast shadow or indirect illumination. The second included shadowing between real and virtual objects enabled but no indirect illumination. The third rendering mode included inter-object shadowing and indirect illumination, causing color bleeding. The study followed a within subjects design and the conditions were administered according to a latin square to minimize the risk of carry-over effects. After the participants had finished all five tasks they were interviewed.

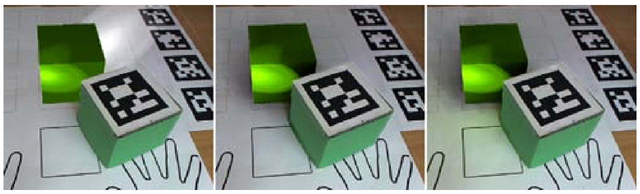

Figure 4. The three different rendering modes (left to right): no shadows/no indirect illumination, shadows/no indirect illumination and shadows/indirect illumination

6.3 Results & Discussion

Twenty one people participated in the study, 15 male and 6 female participants between the age of 19 to 59. All participants but one, who had to be excluded because of color blindness, had normal or corrected to normal eyesight.

It took between 30 and 60 minutes for each participant to finish all five tasks and interview. SPSS 19 was used to analyze the data. Because not all data did meet the requirements for a repeated measures ANOVA (normality, shericity) we analyzed the data using non-parametric Friedman tests.

Our analysis did not show any evidence that the different rendering modes had an effect on task performance. This goes in line with the experiments performed by Thompson et al. [11]. However we have to be cautious in comparing these two experiments since they differ from their setup. In our user-study, the participants had to judge distances less than one meter, whereas in Thompson's experiment the distances ranged from 5 to 15 meters and was based on locomotion. Furthermore they used an immersive VR system whereas we used an AR framework. Therefore we plan to confirm our results and extend our findings through further experiments (see Section 7).

However, we gained other interesting insights. When we designed the tasks we were thinking about disabling occlusion, so that it could not be used as a visual cue. With no occlusion the virtual cube would always be rendered on top of any real-world object - even in situations in which it should be occluded by a real cube. However, since it is a more realistic setup, we decided to allow occlusion. As expected, our study shows that most of the participants used the occlusion cue to place the cubes at the right spot, regardless whether the virtual or the real cubes where manipulated (task 2 & 3). Seven participants recognized the shadows but only one recognized indirect illumination.

In task one the virtual cube was randomly positioned along the main axes and six participants mentioned that it was much easier to estimate the distance on the x and y axis rather than in depth direction. Although there is no significant effect that relates to the comments of the participants the distance estimation error was slightly less for x and y axis. Furthermore the time used for distance estimation is smaller when no shadows and no indirect illumination are calculated. This could indicate that the cognitive load is larger due to more visual cues. However, both effects are not significant and descriptively very small so they could also be just by chance.

In task two the real cube was moved to match the position of the virtual cube. Interestingly, seven participants found task three, manipulating the virtual cube to match the real cube using a computer keyboard, more intuitive and easier. The difference between the two tasks was that the target cube position in task 2 varied on three axes (x, y and z axis) whereas in task 3 it varied only two (x and z axis) but not in height (y axis). Furthermore, in task 3 the participants did not have to change the orientation of the cubes since they were already aligned correctly.

In task 4 and 5 some participants complained that the cube was too large to grab and that the marker for hand tracking disturbed the grabbing process.

We could observe that the participants had rather different ways to approach the tasks. As mentioned in Section 6.2 they had to perform Tasks 2 – 5 as fast and as accurate as possible. Some of the participants focused on speed, others more on accuracy. Some participants made excessive usage of being allowed to move their head to get different viewing angles, while others nearly did not move at all. These different methods probably influenced the final results and therefore should be avoided in future experiments.

7 Future Work

The presented research framework is work in progress and not all envisioned features are implemented. One of these features is stereo rendering. Since the rendering method already pushes the limits of the graphics hardware, rendering a complete second frame is not possible yet while maintaining useable frame-rates. However, many parts in the image pairs are the same and maybe a more sophisticated method can keep the additional rendering overhead quite small.

It is important that the fish-eye lens camera captures an HDR environment map. Our system currently uses scaled LDR environment maps since we do not have the appropriate hardware yet. The calibration process to use an optical tracking system in an augmented scene is crucial. It would be very convenient if there were utility functions available that perform the necessary steps automatically and make calibration easier.

Finally, we want to perform further experiments with altered tasks. New tasks could be similar to the existing ones but with disabled occlusion. The results could be directly compared to the previous ones. Another idea is that the participants have to place the cube on top of the other cube, instead of placing it at the same position. This way the influence of the occlusion cue could be reduced. To get a better experiment setup we want to reduce the size of the cube and use a chin rest to limit the head-movement. Once stereo rendering is supported, we can create tasks that analyze the influence of indirect illumination in a stereo setup.

8 Conclusion

We started this paper by describing what photorealistic augmented reality is and where it can be used. We discussed the current issues that need to be solved and based on that proposed a new research framework to perform perceptual experiments. To our knowledge this is the first research framework that can take real-time global illumination - and dynamic surrounding illumination effects into account. To test the research framework a small preliminary user-study was performed to investigate the influence of different rendering modes on five different tasks. The results indicated that there were no significant effects of these rendering conditions on task performance. However, we plan to perform further experiments to confirm these results with altered tasks as described in the future work section.

References

[1] E. Kruijff, J.E.S. II, and S. Feiner, "Perceptual Issues in Augmented Reality Revisited," *In Proceedings of the 9th IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2010.

[2] G.S. Hubona, P.N. Wheeler, G.W. Shirah, and M. Brandt, "The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization," *ACM Transactions on Computer-Human Interaction*, vol. 6, Sep. 1999, pp. 214-242.

[3] N. Sugano, H. Kato, and K. Tachibana, "The effects of shadow representation of virtual objects in augmented reality," *Mixed and Augmented Reality, 2003. Proceedings. The Second IEEE and ACM International Symposium on*, IEEE, 2003, p. 76–83.

[4] C. Madison, D. Kersten, W. Thompson, P. Shirley, and B. Smits, "The Use of Subtle Illumination Cues for Human Judgement of Spatial Layout," *University of Utah Technical Report UUCS-99-001. Computer Science Dept., Univ. of Utah*, 1999.

[5] H.H. Hu, A.A. Gooch, S.H. Creem-Regehr, and W.B. Thompson, "Visual Cues for Perceiving Distances from Objects to Surfaces," *Presence: Teleoperators and Virtual Environments*, vol. 11, Dec. 2002, pp. 652-664.

[6] G. Nakano, "Generating perceptually-correct shadows for mixed reality," *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, Sep. 2008, pp. 173-174.

[7] J. Lopez-Moreno, V. Sundstedt, F. Sangorrin, and D. Gutierrez, "Measuring the perception of light inconsistencies," *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, ACM, 2010, p. 25–32.

[8] M. Knecht, C. Traxler, O. Mattausch, W. Purgathofer, and M. Wimmer, "Differential Instant Radiosity for mixed reality," *Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on*, IEEE, 2010, p. 99–107.

[9] A. Keller, "Instant radiosity," *Proceeding SIGGRAPH '97 Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, Citeseer, 1997.

[10] P. Debevec, "Rendering Synthetic Objects into Real Scenes: Bridging Traditional and Image-based Graphics with Global Illumination and High Dynamic Range Photography," *25th annual conference on Computer graphics and interactive techniques*, ACM, 1998.

[11] W.B. Thompson, P. Willemsen, A.A. Gooch, S.H. Creem-Regehr, J.M. Loomis, and A.C. Beall, "Does the quality of the computer graphics matter when judging distances in visually immersive environments?," *Presence: Teleoperators & Virtual Environments*, vol. 13, 2004, p. 560–571.

[12] J. Ferwerda, "Three varieties of realism in computer graphics," *Proceedings SPIE Human Vision and Electronic*, 2003, pp. 290-297.

[13] T.J. Hattenberger, M.D. Fairchild, G.M. Johnson, and C. Salvaggio, "A psychophysical investigation of global illumination algorithms used in augmented reality," *ACM Transactions on Applied Perception*, vol. 6, Feb. 2009, pp. 1-22.

[14] M. Elhelw, M. Nicolaou, A. Chung, G.-Z. Yang, and M.S. Atkins, "A gaze-based study for investigating the perception of visual realism in simulated scenes," *ACM Transactions on Applied Perception*, vol. 5, Jan. 2008, pp. 1-20.

[15] P. Rademacher, J. Lengyel, E. Cutrell, and T. Whitted, "Measuring the perception of visual realism in images," *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, Citeseer, 2001, p. 235–248.

[16] D. Drascic and P. Milgram, "Perceptual issues in augmented reality," *PROCEEDINGS-SPIE THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING*, Citeseer, 1996, p. 123–134.

[17] J.E. Cutting, "How the eye measures reality and virtual reality," *Behavior Research Methods Instruments and Computers*, vol. 29, 1997, p. 27–36.

[18] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics*, vol. 21, Jul. 2002.

[19] G. Klein and D. Murray, "Compositing for small cameras," *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, Sep. 2008, pp. 57-60.

[20] J. Fischer, D. Bartz, and W. Strasser, "Enhanced visual realism by incorporating camera image effects," *2006 IEEE/ACM International Symposium on Mixed and Augmented Reality*, Oct. 2006, pp. 205-208.

[21] H. Landis, "Production-ready global illumination," *Siggraph Course Notes*, vol. 16, 2002, p. 2002.

[22] G. Klein and D. Murray, "Parallel Tracking and Mapping for Small AR Workspaces," *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, Nov. 2007, pp. 1-10.

[23] G. Klein and T. Drummond, "Sensor Fusion and Occlusion Refinement for Tablet-Based AR," *Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, pp. 38-47.