Eurographics Symposium on Virtual Environments (2006)
Roger Hubbold and Ming Lin (Editors)

Measuring the Discernability of Virtual Objects in Conventional and Stylized Augmented Reality

Jan Fischer[†] [1] Douglas Cunningham [1] Dirk Bartz[1] Christian Wallraven[2] Heinrich Bülthoff[2] Wolfgang Straßer[1]

[1] WSI / GRIS - VCM,
University of Tübingen,
Tübingen, Germany

[2] Cognitive and Computational Psychophysics,
MPI Biological Cybernetics,
Tübingen, Germany

Abstract

In augmented reality, virtual graphical objects are overlaid over the real environment of the observer. Conventional augmented reality systems normally use standard real-time rendering methods for generating the graphical representations of virtual objects. These renderings contain the typical artifacts of computer generated graphics, e.g., aliasing caused by the rasterization process and unrealistic, manually configured illumination models. Due to these artifacts, virtual objects look artifical and can easily be distinguished from the real environment.

A different approach to generating augmented reality images is the basis of stylized augmented reality [FBS05c]. Here, similar types of artistic or illustrative stylization are applied to the virtual objects and the camera image of the real enviroment. Therefore, real and virtual image elements look significantly more similar and are less distinguishable from each other.

In this paper, we present the results of a psychophysical study on the effectiveness of stylized augmented reality. In this study, a number of participants were asked to decide whether objects shown in images of augmented reality scenes are virtual or real. Conventionally rendered as well as stylized augmented reality images and short video clips were presented to the participants. The correctness of the participants' responses and their reaction times were recorded. The results of our study show that an equalized level of realism is achieved by using stylized augmented reality, i.e., that it is significantly more difficult to distinguish virtual objects from real objects.

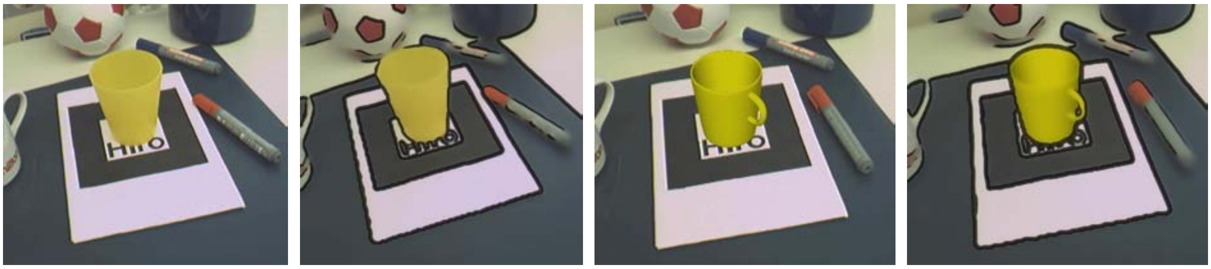
Categories and Subject Descriptors (according to ACM CCS): H.5.1 [Information Interfaces and Presentation]: Artificial, augmented, and virtual realities; I.3.3 [Computer Graphics]: Display algorithms

1. Introduction

Augmented reality (AR) has become a widespread method for enriching the user's environment with virtual objects [ABB*01]. In video see-through augmented reality, a digital video camera continually acquires images of the real surroundings. Graphical objects are then drawn over the camera image, which is displayed as a background image plane. In order to achieve a correct spatial positioning and orientation when rendering the virtual objects, tracking techniques like vision-based marker tracking are normally used [KB99].

In conventional augmented reality systems, the graphical objects are rendered over the camera image using standard real-time graphics algorithms. Low level software libraries like OpenGL or high level scene graphs based on them are often used for this task. The underlying real-time rasterization methods rely on simplified assumptions for illumination and shading. Manually placed virtual light sources are used for the lighting calculations. Simple interpolation methods like Gouraud shading then spread the computed brightness values over the graphical models. The resulting renderings tend to look artificial, and they stand out from the camera image. Even if more sophisticated rendering methods with advanced illumination and shading are used, the problem of mismatched scene generation parameters still persists. Since light sources and material properties are defined during the definition of the AR scene, they generally do not correspond well to the lighting conditions in the real environment.

† e-mail: fischer@gris.uni-tuebingen.de



(a) Real cup in conventional AR (b) Real cup in stylized AR (c) Virtual cup model in conventional AR (d) Virtual cup model in stylized AR

Figure 1: *Two example objects used in the study: One real cup (Fig. 1(a) and 1(b)) and one virtual cup model (Fig.1(c) and 1(d)). All test objects in the study are located directly over the marker used for camera tracking. In order to provide some visual reference for the participants, several background objects are placed near the marker. For each test object, conventionally rendered as well as stylized AR images were recorded and presented to the participants.*

Stylized augmented reality is a different approach to generating augmented video streams [FBS05c]. It attempts to create similar levels of realism in both the camera image and the graphical objects with the help of artistic or illustrative rendering and image filtering. For the psychophysical study described in this paper, a cartoon-like stylization algorithm was used. This method produces augmented images composed of mostly uniformly colored regions, which are enclosed with black silhouette lines. The stylization algorithm is a specialized postprocessing filter, which is applied to the augmented image after the overlay of virtual objects [FBS05b]. The method has been implemented using vertex and fragment shaders running on the programmable graphics processing units (GPUs) of recent graphics cards. Since the same type of stylization is applied to both the camera image and the virtual models, they look more similar than in conventional AR.

In this paper, the results of a psychophysical study measuring the effectiveness of stylized augmented reality are presented. We assume that one evidence for the effectiveness of stylized AR is the degree of difficulty of distinguishing real objects from virtual models. This means that in stylized augmented reality, it should be more difficult for the user to tell whether an object visible in the augmented image is virtual or not. This concept is illustrated in Figure 1: Figure 1(a) shows a real cup in a conventionally rendered image. In this image - and even more so in an interactive real-time AR setup - it is relatively easy to identify the cup as an actual physical object. Figure 1(b), by contrast, contains a stylized image of the same real cup. In Figure 1(d), the stylized augmented reality rendering of a similarly shaped virtual cup, which has been geometrically modelled, is shown. In the latter two, it should be more difficult for an observer to tell if the central object is real or virtual.

For the sake of clarity, we will use the following terminology in this paper. Actually existing items in the environment of the user which are visible in the camera image are called "physical objects" (e.g., see Fig. 1(a) and 1(b)). The term "virtual objects" is used for computer generated graphical models in the augmented image (see Fig. 1(c) and 1(d)). We refer to this distinction between physical and virtual objects as *object type*. As a concept which is orthogonal to these two types of object, we also distinguish between two types of augmented reality rendering. Both physical and virtual objects can be displayed in "stylized" (see Fig. 1(b) and 1(d)) and "conventionally rendered" (see Fig. 1(a) and 1(c)) augmented reality images. The difference between conventionally rendered and stylized is called *AR rendering style*.

The remainder of this paper is structured as follows. Section 2 gives an overview of related work. In Section 3, the algorithm used for the stylization of augmented reality images is briefly summarized. The experimental methodology of the psychophysical study described in this paper is discussed in Section 4. Section 5 presents the experimental results of the psychophysical study. Finally, conclusions drawn from the study are discussed in Section 6.

2. Related Work

An approach which is complementary to the concept of applying stylization to AR images is the attempt to improve the realism of virtual objects. This approach is often referred to as *photometric registration*. This way, a better visual correlation to the camera image can also be achieved. Research has been done into methods of analyzing the real illumination conditions in an augmented reality setup. Examples of this approach include the work of Agusanto et al. on analyzing the distribution of real light sources, which is then used for adapting the representation of graphical objects [ALCS03]. In their system, a mirror ball and special camera are used in a specific procedure for determining the lighting conditions in the scene beforehand. An advanced type of photometric registration is the method developed by Okumura et al. for analyzing the blur in the camera image in order to generate a representation of graphical objects which depends on the current focus plane of the digital camera [OKY06].

Our algorithm for generating a stylized augmented video stream is based on a non-photorealistic image filter. Non-photorealistic, artistic and illustrative rendering and image processing have been areas of very active research for many years. Strothotte and Schlechtweg have published a good survey of methods used in the field [SS02]. One example of an algorithm for the cartoon-like stylization of photographs is the work presented by DeCarlo and Santella [DS02]. Their technique uses a combination of color segmentation and edge detection, which partly inspired our approach. However, this method requires several minutes for processing an input image. An algorithm for the semi-automatic conversion of a real video sequence into a cartoon-like video has been presented by Wang et al. [WXSC04]. This method produces results of good visual quality, but it is an offline algorithm and computationally too expensive for real-time applications.

An application of non-photorealistic rendering to virtual environments was presented by Klein et al. [KLK*00]. Here, non-photorealism was only applied to virtual objects, and no video information was included. A method for integrating non-photorealistic rendering into augmented reality has been presented by Haller and Sperl [HS04]. However, this system applies artistic rendering techniques only to the virtual objects, whereas the camera image is displayed in its original, unprocessed form. The concept of stylized augmented reality as well as an algorithm for the cartoon-like stylization of AR images was first described by Fischer et al. [FBS05c]. Later, an improved approach to the cartoon-like rendering of augmented reality video streams was presented [FBS05b]. This method was also used for generating the images shown to the participants of the psychophysical study described in this paper. As a further extension of the concept of stylized AR, a system for rendering AR videos in an artistic brush stroke style was presented [FBS05a]. Haller et al. have developed an algorithm for displaying complete augmented reality scenes including the virtual objects and the real background in an artistic style [HLB05].

The experiment presented in this paper is a psychophysical examination of the effect of stylization on the perception of realism. A number of different techniques have been used to determine the perceptual realism of computer graphics algorithms [LLC03,MR05,WBCB05]. Some experiments rely on indirect measures, examining the “behavioral realism”. In these experiments, performance on a specific task is compared between the real world and a virtual environment. The similarity of the responses produced by the virtual and real scenes is a measure of the behavioral realism of the system. For example, Mania and Robinson [MR05] compared estimates of presence and subjective lighting quality for both real world scenes and several virtual environments. Presence was measured using standard questionnaires and the subjective response to lighting was measured using sematic differentials (e.g., the scene was rated along several dimensions, such as warm versus cold and relaxing versus tense, using a scale of 1 to 79).

Other experiments determine the realism of a scene more directly. Longhurst et al. [LLC03], for example, showed a series of carefully controlled real scenes, photographs of the real scenes, and printed photographs of rendered versions of the real scenes to participants and asked them a variety of questions, including “was the image real?”. By systematically varying scene properties, such as shadows and lighting quality, these experiments can determine the impact of those properties on perceived realism.

3. Stylized Augmented Reality

This section gives an overview of the algorithm used for stylizing the augmented reality images shown in the psychophysical study. (This algorithm was first presented in [FBS05b]). For each frame, a standard augmented reality pipeline first generates an output image containing the camera image with overlaid virtual objects. This original AR frame is rendered using the graphics hardware and resides in its local frame buffer memory. A postprocessing filter is then applied to it, which is executed by the graphics processing unit (GPU). An overview of the approach is shown in Figure 2.

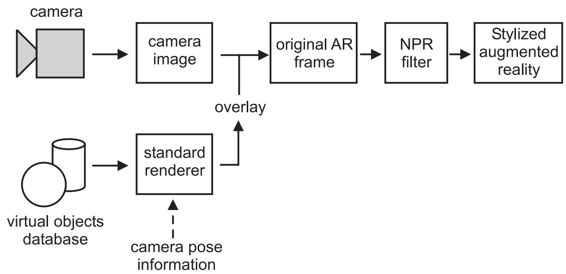


Figure 2: *Overview of the stylized augmented reality pipeline.*

The stylization filter consists of two steps. In the first step, a simplified color image is computed from the original AR frame. The simplified color image is made up of mostly uniformly colored regions generated by a non-linear filter using a photometric weighting of pixels. Several filtering iterations are consecutively applied to the image in order to achieve a sufficiently good color simplification. The second stage of the non-photorealistic filter is an edge detection step executed on the simplified color image. This way, the generated silhouette lines are located between similarly colored regions in the image, which is an approximation of a cartoon-like rendering style. Finally, the simplified color image is combined with the edge detection results by drawing the edge detection responses over it as black lines.

3.1. Generation of Simplified Color Image

A shrunk version of the original AR frame is rendered into the local frame buffer of the graphics card. The non-linear filter is then applied iteratively by using the output image of the last iteration as input texture for the next filtering step. The color simplification filter is inspired by bilateral filtering, which is a widespread method for creating uniformly colored regions in an image [TM98]. The bilateral filter algorithm combines geometric and photometric weights when adding up pixels in the neighborhood of the currently regarded pixel. While the geometric factor gives a greater weight to pixels closer to the current location, the photometric weight suppresses the influence of pixels with very dissimilar color values.

For color segmentation in stylized AR, the only photometric weight is taken into account. Ignoring the geometric weight simplifies the algorithm and reduces its computational complexity. Moreover, this simplified non-linear filter produces good visual results. In addition to disregarding the geometric weight, the filter is also modified so that the photometric weight only depends on the actual color of each pixel. Each pixel is converted into the YUV color space before the filter is applied. In the YUV color space, the Y component represents the brightness of a pixel, while U and V are the chrominance (color) components. For computing the weight of each pixel in the neighborhood, the filter only relies on the U and V coordinates.

We denote the original RGB image function as $\mathbf{f}$ and the corresponding color coordinates in YUV space as $\mathbf{f}_{UV}$. The non-linear filter computes the simplified RGB image $\mathbf{h}$ using the following equation:

$$\mathbf{h}(\mathbf{x}) = k^{-1}(\mathbf{x}) \sum_{\xi \in \Omega_{\mathbf{x}}} \mathbf{f}(\xi)\, s(\mathbf{f}_{UV}(\xi), \mathbf{f}_{UV}(\mathbf{x})) \qquad (1)$$

In Equation 1, $\mathbf{x}$ is the currently regarded point in the output image. A weighted sum is computed over image points ξ in the neighborhood $\Omega_{\mathbf{x}}$ of $\mathbf{x}$ in the input image. A quadratic image area is used as neighborhood for the summation. The weight $s(\mathbf{f}_{UV}(\xi), \mathbf{f}_{UV}(\mathbf{x}))$ depends on the similarity of values in the color channels, $\mathbf{f}_{UV}(\xi) - \mathbf{f}_{UV}(\mathbf{x})$. In our algorithm, s is a Gaussian function. The standard deviation of this Gaussian function determines the properties of the color simplification and is a parameter of the algorithm. In order to maintain the overall brightness of the image, the weighted sum is divided by the normalization factor $k(\mathbf{x})$, which is computed as the sum of all factors $s(\mathbf{f}_{UV}(\xi), \mathbf{f}_{UV}(\mathbf{x}))$ in Equation 1.

The effect of this non-linear filter is that an averaging of pixels only occurs in places where nearby pixels have similar colors. In such places in the image, $s(\mathbf{f}_{UV}(\xi), \mathbf{f}_{UV}(\mathbf{x}))$ is large. If near the currently regarded pixel colors are present which are far away in color space, they are not taken into account. Thus strong edges in the image are preserved. A small local neighborhood of 5 x 5 pixels is used for the weighted summation. Figure 3 shows an example of a simplified color image computed for an augmented reality frame.

(a) Original AR frame

(b) Simplified Color Image

Figure 3: *Generation of the simplified color image for an original AR frame. In the augmented reality scene, a virtual plane model is overlaid over the camera image.*

3.2. Edge Detection Step

After the simplified color image has been generated, the edge detection step is performed. The Sobel edge detection filter is used for computing the partial derivatives of color channel values along the x-axis and the y-axis. Here again, the pixels are converted into YUV color space beforehand. We denote the image function of the simplified color image as $\mathbf{S}$, consisting of the channels (S_Y, S_U, S_V). For each of the color channels, two partial derivatives are calculated. Based on the partial derivatives, gradient magnitudes $(|\nabla S_Y|, |\nabla S_U|, |\nabla S_V|)$ are computed.

An edge detection response is then calculated for each pixel using the gradient magnitudes. This response value is obtained through the weighted averaging of the local contrast in the intensity (Y) and color (U,V) channels. The relative weight of the intensity and color contrasts is determined by the parameter $\alpha \in [0;1]$. Equation 2 shows the computation of the edge detection response for the simplified color image, $edge_{(S)}$. Using this method, the edge detection process can generate responses in locations with homogeneous intensities, where the color channel gradient is large. By adjusting parameter α, intensity contrasts or color contrasts can be emphasized when locating silhouette edges.

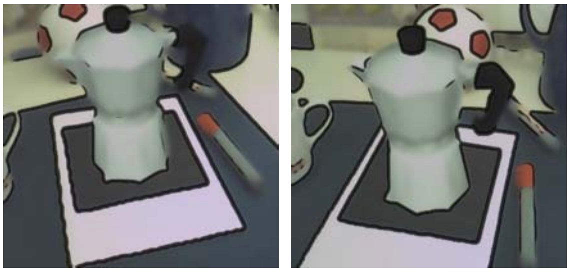
$$edge_{(S)} = (1-\alpha) \cdot |\nabla S_Y| + \alpha \cdot \frac{|\nabla S_U| + |\nabla S_V|}{2} \qquad (2)$$

The edge detection response $edge_{(S)}$ is filtered with the *smoothstep* function. This function is provided by the OpenGL Shading Language, which is used by the implementation of the algorithm. It returns a value of zero for edge detection responses below a threshold s_0, and a value of one for responses above a threshold s_1. Between the two thresholds, smooth Hermite interpolation is used . The values s_0 and s_1 are parameters of the algorithm. They determine the minimum edge detection response necessary for generating a silhouette, and how steeply the silhouette intensity increases. The resulting filtered edge detection response I_o is used for drawing silhouette lines in the output image. Each pixel in the output image is computed as $(1-I_o) \cdot S(x_s, y_s)$, i.e., as the simplified color image pixel multiplied by $(1-I_o)$. This way, greater edge detection responses lead to darker output pixels. Figure 4 shows the final output image generated for the original AR frame in Figure 3(a).

Figure 4: *Final output of the non-photorealistic filter.*

4. Experimental Methodology

The psychophysical study presented in this paper was designed as an offline task. This means that the participants did not wear a head-mounted display driven by an interactive augmented reality application. Optical marker tracking, which we use in our AR system, often fails in an interactive setting with inexperienced users. This leads to the inadvertent disappearance of all virtual objects, revealing that they cannot be physical. Moreover, such a setup would have complicated the execution of the study. Since an assessment of visual differences was the main objective of the study, we only showed recorded still images and short video clips to the participants on a conventional monitor.



(a) First frame of video clip (b) Last frame of video clip

Figure 5: *First and last frame of an AR video clip showing the stylized rendering of a virtual coffeemaker.*

The test scenes constituting the psychophysical study show different virtual and physical objects. For each object, one conventionally rendered and one stylized video clip were recorded by grabbing real-time frame buffer images from the actual augmented reality application. A standardized setting consisting of a large optical marker with some real background objects was used when recording the video clips (see Fig. 1). The currently regarded object is always centered directly over the marker. Each video was shot with a standardized camera path. This camera movement is illustrated in Figure 5. The first frame of each video clip was also used as still image for the static experiments. The recorded video clips and still images were presented to the participants. They were asked to decide whether the displayed object is physical or virtual. The correctness of the response and the participant's reaction time were recorded in a protocol file and later evaluated.

4.1. Outline of the Study

Conventional and stylized presentations of 15 physical and 15 virtual objects were presented to 18 individuals in a psychophysical experiment. The individuals, who participated in return for financial compensation at standard rates, were randomly assigned to one of two groups. One group was presented with a video sequence of the camera moving around the object (the "Dynamic" group). The other group was presented with the first frame of the video sequence (the "Static" group). The participants' task was to determine if the central object in each image was real or virtual. A selection of objects used in the study is shown in Figures 8 and 9 (color plate).

4.2. Stimuli

Using the augmented reality system and real-time stylization algorithm described in Section 3, the 15 real and the 15 virtual objects were recorded using a simple, partially curved camera trajectory (see Figure 5). Each object was filmed both in conventional mode as well as in stylized mode, yielding 60 recordings. Care was taken so that the trajectory of the camera was as identical as possible across the 60 recording sessions. The video sequences were approximately 4 seconds each. During the study, the resulting images were scaled from their original size of 640 x 480 pixels to 1024 x 768 pixels, filling the screen of the 21 inch computer monitor. Since the participants sat at a distance of approximately 0.5 meters from the computer screen, the images subtended approximately 43.6 by 33.4 degrees of visual angle. (Note that while the full image subtended 43.6 by 33.4 degrees, the experimental object subtended a substantially smaller angle.)

4.3. Procedure

The participants were given an instruction sheet describing the experimental task. In particular, the participants were told that they would be presented with several images containing a tracking marker, on top of which would be either a physical object or a virtual object. They were given several example photographs. (The example object was not used during the main experiment.) The participants were also told that sometimes the images would be rendered in conventional manner and sometimes in a stylized manner, and were again given examples.

Each participant saw all 60 trials in a different random order. Each trial began when the participant pressed a key and ended when they entered their answer. For the Dynamic group, the video sequence was shown in a continuous loop with a 250 ms blank screen between repetitions. For the Static group, the first frame of the recording was shown on the screen. The accuracy and speed of the participants' answers were recorded and separately subjected to a repeated measures analysis of variance (ANOVA), with *AR rendering style* (conventional versus stylized) and *object type* (physical versus virtual) as within-subjects factors and *motion type* (static versus dynamic) as a between-subjects factor. The data from one participant in the Static group were not analyzed, as the participant did not follow the instructions.

5. Results

In the following, we will describe the statistical analysis of the experimental results with respect to both recognition accuracy and reaction times.

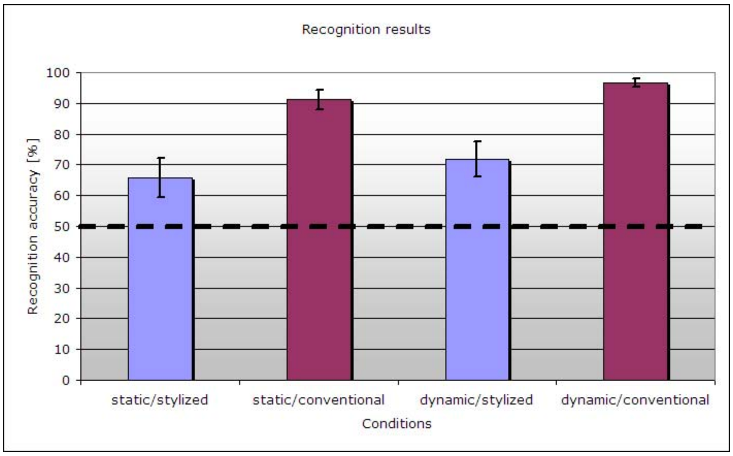
5.1. Accuracy

Participants found it significantly harder to tell the difference between physical and virtual objects in stylized AR than in conventionally rendered images (69% versus 94% accuracy ratings, respectively). This is reflected in the significant main effect for *AR rendering style* (F(1,15)=57.345, p<0.0001). (P-values reflect the probability that the statistical result could have arisen by pure chance; a p-value smaller than 0.05 is usually taken as evidence for a non-random result.) The fact that the overall accuracy for stylized AR images is still significantly above chance, suggests that while the stylization helped to mask the difference between physical and virtual, it did not completely eliminate it. (Chance level accuracy is what one would expect if the participants were blindly guessing, for example, if they were completely unable to tell the difference between physical and virtual objects. For this experiment, which used a two alternative forced-choice task, chance is 50%.) The fact that the accuracy rate is so low, however, clearly demonstrates the general effectiveness of using stylization. Future experiments with either stronger stylization or different stylization algorithms should easily be able to completely mask the difference between physical and virtual objects.

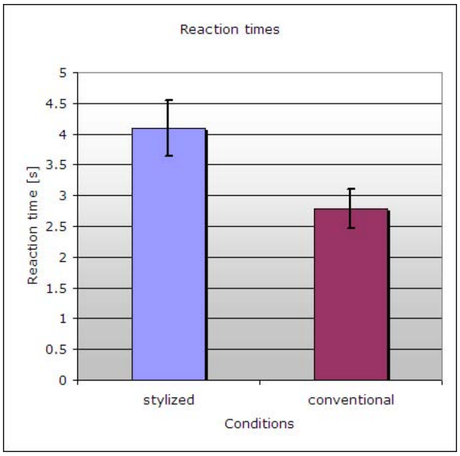
The results are remarkably consistent between the Static and Dynamic groups (see Fig. 6(a)). This is reflected in both the lack of an effect of *motion type* ($F(1,15) = 2.154, p > 0.16$, *not significant*) and the lack of an interaction between *motion type* and *AR rendering style* ($F(1,15) = 0.008$, *not significant*). This clearly shows that stylization for dynamic sequences did not introduce any artefacts that might have resulted in easier detection of virtual versus physical objects. On the contrary, it seemed that the slight jitters introduced by the sometimes imperfect camera tracking were effectively masked out by the stylization.

The main effect for *object type* was close to being significant ($F(1,15) = 3.875, P < 0.068$): Overall, participants had more difficulty identifiying physical objects than identifying virtual objects (77% versus 86% correct, respectively). This difference is driven almost completely by the physical objects shown as stylized AR images: In the conventionally rendered scenes, virtual and physical objects were correctly labeled 92% and 96% of the time, respectively. In the stylized AR images, however, virtual and physical objects were correctly labeled 80% and 58% of the time, respectively. This trend, which is reflected in the significant interaction between *object type* and *AR rendering style* ($F(1,15) = 6.985, P < 0.02$), strongly suggests that most errors are due to the incorrect labeling of physical objects in the stylized images. This result shows that our stylization technique was particularly successful in making physical objects almost indistinguishable from virtual objects.

The effect that virtual objects were relatively more accurately identified compared to physical objects in the stylized AR images is probably caused by the geometric models used in the study. Some of these virtual models are of a rather low graphical quality. Several of the virtual objects have rather unrealistic colors (e.g., almost completely black), and most of them are uniformly colored. Another example of an easily identifiable virtual object is shown in Figure 7. The spout of the teapot shown in these images is incorrectly modeled,



(a) Recognition accuracy in percent correct. The dashed line indicates chance level.



(b) Reaction times in seconds.

Figure 6: *Recognition accuracy and reaction times measured in the study. Error bars in both graphs represent standard error of the mean.*

so that it is displayed as translucent. This was probably a strong hint to most participants that this object is not a physical teapot. Nevertheless, the fact that even in this condition, participants showed a clear decrease in performance testifies to the validity of using stylized AR to blur the distinctions beween the virtual and the real world.

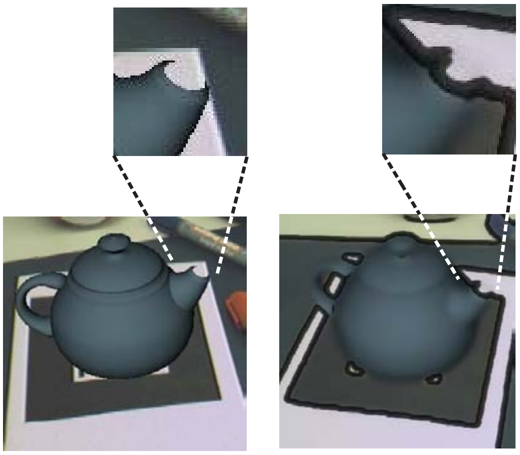


Figure 7: *Graphical error in the virtual model of a teapot, which was used in the study. Due to incorrect backface culling, the spout of the teapot is only partially displayed. This made it easier for participants to identify the teapot as a virtual object, even in stylized AR (right image).*

5.2. Reaction Times

Interestingly, the only statistically significant effect we could find in the reaction times data was that people are slower when assessing stylized AR images than when assessing conventionally rendered AR images (4.0 seconds versus 2.8 seconds, $F(1,15)=17.059$, $p<0.001$; see Fig. 6(b)). No other main effects or interactions reached significance. In particular, we found no difference between the Static and Dynamic conditions showing that dynamic information did not help to speed up the task. Most importantly, virtual and real objects were processed equally fast in all conditions suggesting that participants used the same strategies for both object types. Finally, the difference between stylized and conventionally rendered images reflects the fact that participants had to make a difficult decision in case of stylized images. This, however, does *not* mean that stylization in itself would result in a less effective AR environment, but rather points towards the difficulty of the task.

6. Conclusions

The results of our psychophysical study showed that presenting the scenes in a stylized manner successfully reduced the detectable differences between physical and virtual objects. Critically, the results were nearly identical for the Static and Dynamic groups, suggesting that the results should generalize to an interactive version of the task in which dynamic information should play an even more important role. Finally, the majority of the errors consisted of falsely believing that some of the physical objects presented in stylized AR were actually virtual objects. This highlights the success of the stylization algorithm in generating a consistent representation of virtual and physical image elements.

As mentioned in Section 5, some of the virtual object models used in our study were of a rather low graphical quality. It can be assumed that a study based on better graphical models would yield an even more decreased recognition accuracy.

The problem presented to the participants of the psychophysical study described in this paper was relatively simple. However, while it could be argued that the experimental setup used in the study was very different from most real-life AR systems, the results of the study indicate that stylized AR is effective. It has been shown that in stylized AR images and video sequences, it is significantly more difficult to distinguish physical objects from virtual objects (and vice versa). This means that a novel user experience is created by applying stylization algorithms in AR, and it also points to an improved "immersion" in the augmented scene.

We plan to further examine the effects of stylized AR in future psychophysical studies. These could include the repetition of a similar study as described in this paper in a more life-like AR setting, e.g., by using a head-mounted display for displaying the images and videos. In a later stage, a study could be executed in an interactive, real-time augmented reality system. Finally, more complicated tasks could be designed which have to be performed by participants both in conventional and stylized augmented reality. An example of such an advanced task performance experiment would be the problem of finding a virtual object (or several virtual objects) in a room in conventional and stylized AR.

This work has been supported by project VIRTUE in the focus program on "Medical Navigation and Robotics" (SPP 1124) of the German Research Foundation (DFG).

References

[ABB*01] AZUMA R., BAILLOT Y., BEHRINGER R., FEINER S., JULIER S., MACINTYRE B.: Recent Advances in Augmented Reality. *IEEE Computer Graphics and Applications 21*, 6 (November/December 2001), 34–47.

[ALCS03] AGUSANTO K., LI L., CHUANGUI Z., SING N.: Photorealistic Rendering for Augmented Reality using Environment Illumination. In *Proc. of IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)* (October 2003), pp. 208–216.

[DS02] DECARLO D., SANTELLA A.: Stylization and Abstraction of Photographs. In *Proc. of ACM SIGGRAPH* (July 2002), pp. 769–776.

[FBS05a] FISCHER J., BARTZ D., STRASSER W.: Artistic Reality: Fast Brush Stroke Stylization for Augmented Reality. In *Proc. of ACM Symposium on Virtual Reality Software and Technology (VRST)* (November 2005), pp. 155–158.

[FBS05b] FISCHER J., BARTZ D., STRASSER W.: Reality Tooning: Fast Non-Photorealism for Augmented Video Streams. In *IEEE and ACM ISMAR Poster Session* (October 2005), pp. 186–187.

[FBS05c] FISCHER J., BARTZ D., STRASSER W.: Stylized Augmented Reality for Improved Immersion. In *Proc. of IEEE Virtual Reality* (March 2005), pp. 195–202.

[HLB05] HALLER M., LANDERL F., BILLINGHURST M.: A Loose and Sketchy Approach in a Mediated Reality Environment. In *Proc. of ACM International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia (Graphite)* (December 2005).

[HS04] HALLER M., SPERL D.: Real-Time Painterly Rendering for MR Applications. In *Proc. of ACM International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia (Graphite)* (June 2004), pp. 30–38.

[KB99] KATO H., BILLINGHURST M.: Marker Tracking and HMD Calibration for a video-based Augmented Reality Conferencing System. In *Proc. of IEEE and ACM International Workshop on Augmented Reality (IWAR)* (October 1999), pp. 85–94.

[KLK*00] KLEIN A., LI W., KAZHDAN M., CORREA W., FINKELSTEIN A., FUNKHOUSER T.: Non-Photorealistic Virtual Environments. In *Proc. of ACM SIGGRAPH* (July 2000), pp. 527–534.

[LLC03] LONGHURST P., LEDDA P., CHALMERS A.: Psychophysically Based Artistic Techniques for Increased Perceived Realism of Virtual Environments. In *Proc. of the ACM International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa (Afrigraph)* (2003), pp. 123–132.

[MR05] MANIA K., ROBINSON A.: An Experimental Exploration of the Relationship between Subjective Impressions of Illumination and Physical Fidelity. *Computers & Graphics 29* (2005), 49–56.

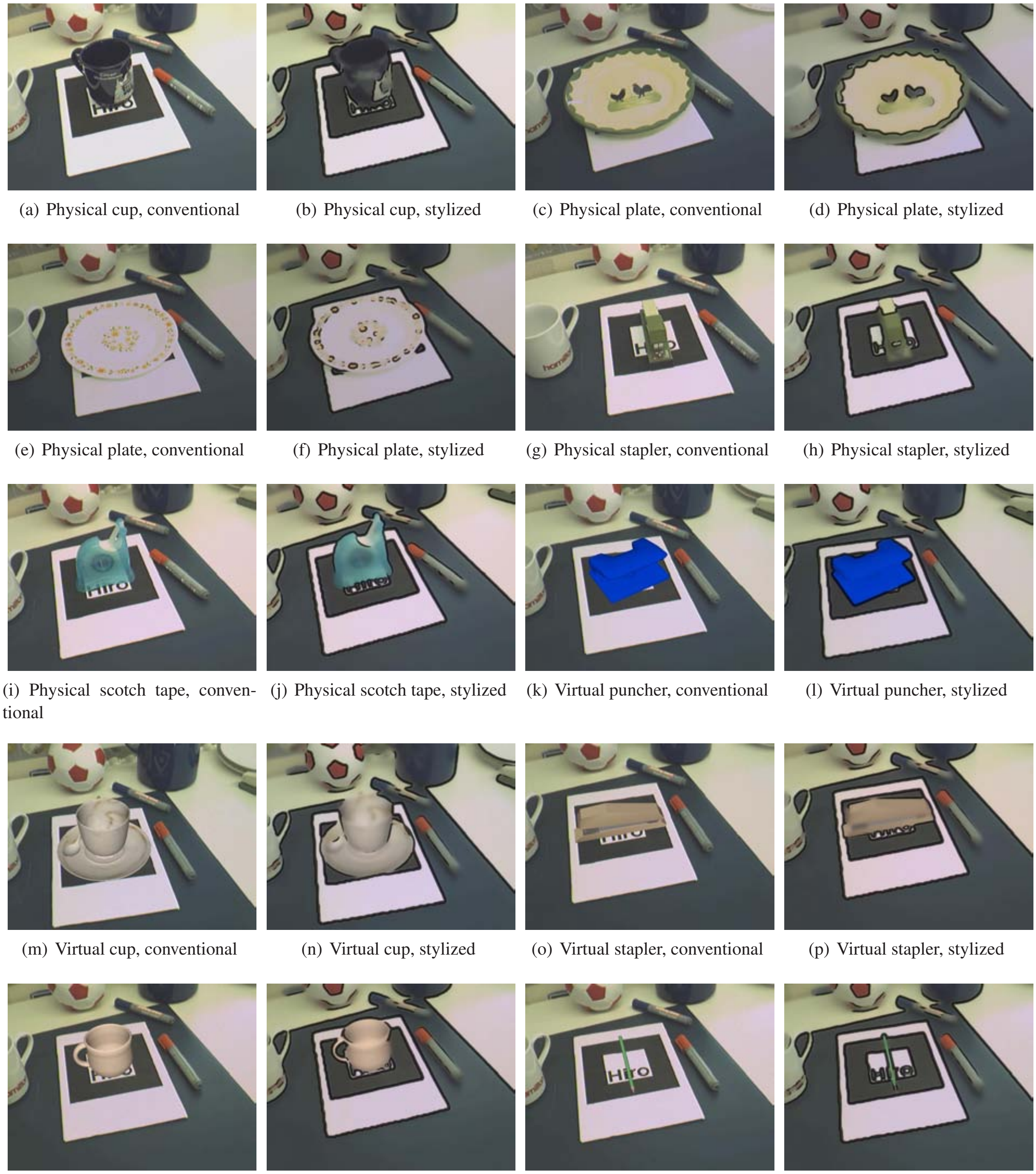
[OKY06] OKUMURA B., KANBARA M., YOKOYA N.: Image Composition Based on Blur Estimation from Captured Image for Augmented Reality. In *Proc. of IEEE Virtual Reality* (2006).

[SS02] STROTHOTTE T., SCHLECHTWEG S.: *Non-Photorealistic Computer Graphics - Modelling, Rendering, and Animation*. Morgan Kaufmann Publishers, 2002.

[TM98] TOMASI C., MANDUCHI R.: Bilateral Filtering for Gray and Color Images. In *Proc. of IEEE International Conference on Computer Vision (ICCV)* (1998), pp. 839–846.

[WBCB05] WALLRAVEN C., BREIDT M., CUNNINGHAM D., BÜLTHOFF H.: Psychophysical Evaluation of Animated Facial Expressions. In *Proc. of the Symposium on Applied Perception in Graphics and Visualization (APGV)* (2005), pp. 17–24.

[WXSC04] WANG J., XU Y., SHUM H., COHEN M.: Video Tooning. In *Proc. of ACM SIGGRAPH* (August 2004), pp. 574–583.



(a) Physical cup, conventional (b) Physical cup, stylized (c) Physical plate, conventional (d) Physical plate, stylized

(e) Physical plate, conventional (f) Physical plate, stylized (g) Physical stapler, conventional (h) Physical stapler, stylized

(i) Physical scotch tape, conventional (j) Physical scotch tape, stylized (k) Virtual puncher, conventional (l) Virtual puncher, stylized

(m) Virtual cup, conventional (n) Virtual cup, stylized (o) Virtual stapler, conventional (p) Virtual stapler, stylized

(q) Virtual teacup, conventional (r) Virtual teacup, stylized (s) Virtual pencil, conventional (t) Virtual pencil, stylized

Figure 8: *Some of the objects shown in the psychophysical study.*