Helena Pongrac
Human Factors Institute
Universität der Bundeswehr München
D-85577 Neubiberg, Germany

Jan Leupold*
Stephan Behrendt
Institute for Real-Time Computer Systems
Technische Universität München
D-80333 Munich, Germany

Berthold Färber
Human Factors Institute
Universität der Bundeswehr München
D-85577 Neubiberg, Germany

Georg Färber
Institute for Real-Time Computer Systems
Technische Universität München
D-80333 Munich, Germany

Human Factors for Enhancing Live Video Streams with Virtual Reality: Performance, Situation Awareness, and Feeling of Telepresence

Abstract

Telepresence systems should be designed to assist the human operator as much as possible to fulfill his task. In order to support the user concerning the visual modality, a system was designed that presents virtual reality images combined together with camera images captured at the remote teleoperator environment.

In this work, two experiments were conducted. In the first, it was shown that presenting a widened field of view to the human operator enhances the human performance and the feeling of telepresence. In the second, it was examined how the transition between video and virtual views has to be designed. Relevant criteria of this transition were chosen and the results show that the operator's rating of quality, feeling of telepresence, and situation awareness are positively affected by variations of the transition parameters. Furthermore, a trade-off between the rating of quality and the situation awareness was observed.

A parameter selection scheme is presented which can serve as a design guideline for combining video and virtual views depending on the desired application.

1 Introduction

In most telepresence systems, a visual representation of the remote teleoperator scene is presented to the human operator. A standard device for feeding back the visual information is the head-mounted display (HMD). Images captured by cameras at the remote scene are transmitted to the operator side and displayed on this HMD. Thus only observations from a set of concrete positions are available. A second drawback of this scenario is that the field of view of these cameras is inherently limited by their optics.

A predictive display was developed in a previous research project (Burkert et al., 2004), which can be used to widen the displayed field of view by combining the camera images with a photorealistic virtual model (Leupold & Behrendt, 2005). This combination has been termed *hybrid view* or *hybrid display*.

Presence, Vol. 16, No. 5, October 2007, 488–508

*Correspondence to leupold@rcs.ei.tum.de

It has been shown that it is reasonable to combine real video views with a virtual model as in augmented reality applications to visualize obscured or covert objects (Furmanski et al., 2002; Dai, 2001). In so doing, the human performance relating to displaying the real remote scene is improved in areas such as better reaction times in telesurgical applications (Chios & Linney, 2004) or a greater amount of remembered objects (Macchiarella & Vincenzi, 2004).

Furthermore, there exists a positive correlation between human performance and the feeling of telepresence as, for example, in spatial searching tasks (Nash et al., 2000). The feeling of telepresence constitutes the "sense of being there" or a "perceptual illusion of nonmediation" in regard to telepresence applications (Insko, 2003; Nash et al., 2000). This means that the smaller the perceived difference between the presented and the remote environment, the higher the feeling of telepresence. Therefore, the feeling of telepresence represents an important construct that has to be considered in telepresence systems.

The general goals of this work are first to show the benefits of the hybrid display compared to camera images only, and second to find an optimal transition between the camera images and the photorealistic virtual model.

After a system description in the following section, more information on the focused concepts is given in Sections 3 and 4. In the remainder of this article, two experiments and the associated results are reported in detail.

2 System Overview

The system was designed with a typical telepresence scenario in mind: a robot in a remote environment controlled by a human operator. In order to present the remote environment to the operator, cameras are mounted on a pan-tilt unit on the robot.

The operator wears a HMD (Sony LDI-D100BE) which is tracked in 6 degrees of freedom (Ascension Flock of Birds Tracker). The rotation angles of the pan-tilt unit can be set to follow the tracked HMD rotation. The images captured by the cameras are sent to the operator side as a *live video stream*. The field of view is restricted through the camera's optics to 37.9° horizontally and 28.7° vertically. This is the maximum presentable field of view using only camera images, no matter which kind of hardware display is used to present the live video stream. As already mentioned in the Introduction, the *hybrid display* allows a widening of the presented field of view using virtual reality: The received live video stream is presented (*video view*) together with a texture-mapped 3D representation of the remote environment. This 3D representation is used to generate photorealistic *virtual views* that can be used to fill the peripheral visual field that is not covered by the video view. A drawback of the virtual view is the time needed to update the 3D representation and the corresponding textures. Therefore a virtual view alone should not be used for telemanipulation tasks and thus the operator should always be aware which parts are virtual views. To give the operator a cue (*view marking*), augmentation of the virtual view can be used.

A drawback of the video view is that no numerical depth information is transmitted in the stream, meaning that the 2D image information is drawn in a 3D environment. If there is no difference between the operator observation position and the camera position (with respect to translation), no depth information is needed for a correctly presented hybrid view. Otherwise an error in the generated view is unavoidable. This error is called *reprojection error* (Leupold & Behrendt, 2005).

The reprojection error describes the displacement of video pixels in regard to their correct position on the display screen. This displacement is due to missing numerical depth information and is caused by a translational difference between the camera and the observer positions. This translational difference is reported when referring to the reprojection error in the remainder of this article. A minimal error is unavoidable due to the precision of camera calibration. This is apparent in the area of transition between virtual and video view in particular. The reprojection error represents a perceptual distortion which signifies a violation of the perspective of a presented scene (for further details see Leupold & Behrendt, 2005).

The design of the spatial transition, that is, the algorithm for drawing the border regions of the video view, can be implemented in many ways. Two algorithms were used in the experiments, which both paint the video view over the virtual view. The first algorithm simply replaces each pixel with the video view (no blending), while the second combines the color values of both views in the border regions (blending, see Figure 3; later in the article). The following three parameters do not have a technical optimum, but depend on human perception: maximum allowed translational difference (reprojection error), type of view marking, and blending strategy. They are subject to Experiment 2 in Section 6.

3 Effects of Restricting the Field of View

In telepresence applications where HMDs are used, it has been found that human performance is affected in regard to distance estimation, searching tasks, and detection tasks, as well as the sensation of comfort. There is evidence that the impacted performance is due to the restricted field of view (Alfano & Michel, 1990; Semmlow et al., 1990; Low et al., 2001; Nash et al., 2000).

In previous studies, it has been shown that the human performance was ameliorated in virtual and augmented realities using HMDs when visual information had been presented in the peripheral visual field, or when the subjects had been allowed to move their head (Willemsen & Gooch, 2002; Creem-Regehr et al., 2003; Witmer & Kline, 1998; Jay & Hubbold, 2003).

Consistent with this finding is the conclusion of Psotka et al. (1998), who stated that human operators use their natural field of view (120° vertically, 180°–200° horizontally) as the primary standard for the estimation of distances and the perception of directions and one's own position or movements.

Peripheral awareness, that is, the display of the environment in the peripheral field of view, turned out to be beneficial even when the peripheral view was not necessary for performing a given task such as searching for objects or object detection (Low et al., 2003; Ilie et al., 2004; Waller, 1999).

It can be concluded that peripheral awareness leads to a reduction of the visual burden. So, more cognitive capacity is available for performing other cognitive or visual tasks such as searching or reaching for objects in virtual realities.

Hitherto, it was assumed that a widened field of view leads to a higher feeling of telepresence (Tang et al., 2005; Rosenblum & Macedonia, 2005).

Section 5 describes an experiment and the associated results where a restricted and a widened field of view are tested against one another in order to determine the positive effects caused by the widened field of view of the hybrid display (Experiment 1).

4 Designing the Combination of Real and Virtual Views

As already mentioned in Section 2, the design of the combination of video and virtual views has a major impact on the resulting representation.

First of all, the parameters that determinantly constitute this combination had to be selected, as well as the variables that have to be gathered from the subjects.

It was decided to select the view marking, the transition, and the reprojection error because these independent variables seemed to be the most important ones to evaluate.

The marking of which display variant is seen was operationalized through visual augmentation of the virtual view. Hitherto the augmentation of virtual views have either not been conceptualized because the visual augmented objects appeared sufficiently artificial (Paloc et al., 2004) or the video view was superimposed by the virtual view without transition (Ellis & Menges, 1998; Thompson et al., 2001; Hua et al., 2004; Nakatsuru et al., 2003). However, this superposition can serve as a distractor that impacts human performance (Counsell, 2004). Others displayed real and virtual views of a scene in parallel without transition, either using different displays (Dai, 2001) or augmenting additional pictures or pictures with text (Ou et al., 2004; Bell et al., 2002; Azuma & Furmanski, 2003; Bell & Feiner, 2000).

Regarding the transition, an attempt to realize an optimal blending was made, arguing that this has a positive effect on the feeling of telepresence, but without examining this assumption directly (Low et al., 2001; Raskar et al., 1999; De Bruyn, 1997; Reitsma & Pollard, 2003; Livingston, 2005).

No systematic research has so far been carried out to investigate the effects of the reprojection error on human performance and the feeling of telepresence (Ilie et al., 2004).

To the authors' knowledge, no studies have been conducted to investigate which design of the combination has to be adopted for diverse applications.

On the human operator's side, it is important that the quality of the combination and that the feeling of telepresence are sufficiently highly rated because the human has to sense at least a minimal amount of comfort as a necessary condition for performance to occur (Nash et al., 2000).

Moreover, the human operator has to be aware of which view variant he or she is seeing to judge the reliability of the presented image data due to uncertainties of the virtual view. These uncertainties are caused by limited update-rates of the virtual model or by the inaccuracy of the remote sensors (Dai, 2001; Azuma, 1997). Thus, the situation awareness is an important factor for ensuring safe teleoperation. This psychological construct includes the perception of the (relevant) elements in a dynamic environment in a given situation, the understanding of their meaning, and the prediction of their following states (Endsley, 2000).

In Section 6 an experiment together with its associated results is reported that was conducted to evaluate which combination should be used for different applications (Experiment 2).

5 Experiment 1

In order to determine the positive effects of a widened field of view shown by the hybrid display, as opposed to the video view only, an experiment was conducted whereby the human performance displaying the combined camera and virtual views was compared with the performance displaying the video view only.

Several performance measures have been realized (see Section 5.4).

The literature review (Section 3) led to the following expectations.

- When displaying the hybrid view, that is, a widened field of view, there should be a greater attention to detail, which means that a greater number of small objects should be identified than with video view only.
- The estimated distance of objects should be ameliorated with the hybrid view compared to video view only.
- The subjects' feeling of telepresence should be better in the condition with the hybrid display than in the condition with video view only.
- The accuracy of the recalled position of objects in a presented scene should be better when the combined camera and virtual views are displayed than when only the video view is displayed.

5.1 Technical Setup

In the setup for this experiment, the operator wears a HMD. On the screen of the HMD either only the received live video stream or the hybrid view were displayed (see above). The operator's head position was tracked and the rotation component was sent to the remote environment to control the rotation angles of the pan-tilt unit. The speed of the camera following was set to 0.5 rad/s.

On the remote side, each frame captured by the camera was sent together with position information (rotation and translation) to the operator side at a rate of 33 frames per second (video size: 640 × 480 pixels). The tracked HMD position was also used as an input variable for the hybrid/video view generation running at 50 Hz. Due to communication delays and the limited rotation speed of the pan-tilt unit, the positions of the received frames of the live video stream do not, in general, coincide with the current operator head position. The

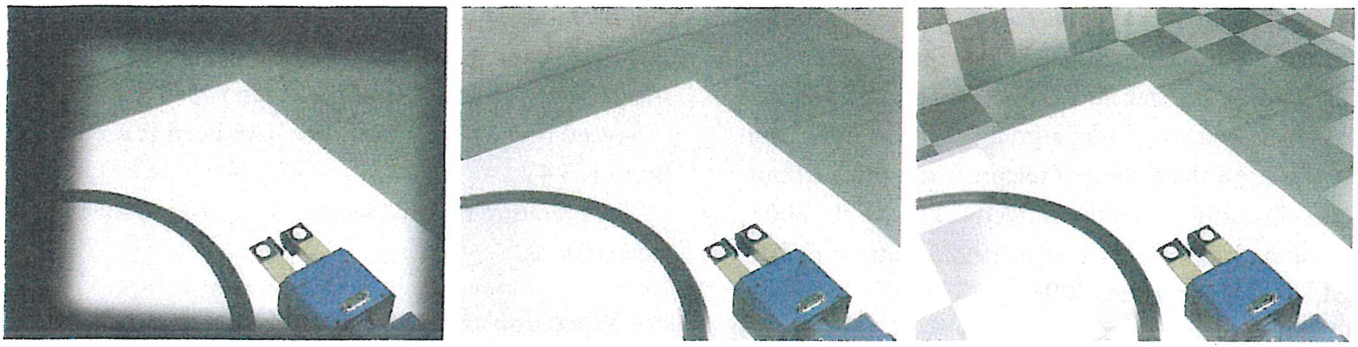


Figure 1 *Video view, hybrid view, and augmented hybrid view with a checker pattern (from left to right).*

display algorithm needs to take this into account while generating the view. In Experiment 1, the operator's head translation component was implicitly set as identical to the translation component of the last video frame received to avoid reprojection errors here. The generated view always reflected the operator's current head rotation, with visual information from the remote side displayed at the position from which it was captured. The horizontal field of view of the generated view was 41.4°, while the field of view of the camera was 37.9° (see Figure 1). While the difference seems to be negligible in the static case, it has to be considered that the camera rotation speed was limited so during quick head rotations a larger part of the virtual view was presented.

The 3D representation of the remote environment was loaded and textured before the experiment and was not updated during it. Four small objects, namely a small champagne bottle, a tennis ball, a green box and a No Smoking sign, were details of the scene. Two of the objects were included in the 3D representation, either as textured polygonal data (the green box) or only as texture (the No Smoking sign). These objects were seen in both views. The other two objects were only present in the video view. This was done to simulate real applications, where the virtual model may not always be brought up-to-date to the newest configurations.

The left and middle images in Figure 1 show the experimental conditions video view and hybrid view, respectively. The right picture illustrates the virtual part of the hybrid view by augmentation with a checker pattern.

5.2 Subjects

Ten female and 27 male subjects participated. Their age range was 21–67 (M_{age} = 32). They all had normal vision or corrected to normal vision. Fourteen of them wore no glasses or contact lenses at all, 5 reported wearing glasses only for special circumstances such as driving, and 18 declared that they wear glasses or contact lenses permanently. Most of the subjects (n = 35) had little or no previous experience with HMDs, only 2 of the participants had had more experience with HMDs. The average time spent playing 3D computer games was stated to be 2 hours a month over all subjects.

Before and during the experiments, none of the subjects was given information about the intentions or hypotheses of the experiments. Afterward, they received a full explanation. Their personal data were recorded anonymously.

5.3 Procedure

To begin with, the subjects were shown a training scene through which they could become familiar with the HMD and the system. This scene was also used for individually adjusting the 3D-stereo projection for each participant.

When the subjects became acquainted with the system, the experimental scene was presented to the subjects for 25 s through an HMD, either with or without the additional virtual model displayed. The two conditions were balanced across subjects, so 19 subjects participated in the condition where the virtual model combined with the video view was displayed, whereas 18 subjects were in the condition with video view only.

After being presented with this scene the subjects were requested to rank how strongly they felt immersed in the shown environment on a scale of 1 to 7, where 1 meant "very weakly" and 7 "very strongly," in order to determine their feeling of telepresence.

Then it was asked which of the four objects mentioned above (bottle, green box, tennis ball and No Smoking sign) the subjects had identified while looking at the scene. For every identified object, the denoted estimated ego-distance was recorded. In addition, the subjects were shown a screenshot that did not contain these four objects, on which they were required to mark the position of the identified objects. Following this procedure, the subject proceeded to carry out Experiment 2 (see below).

At the end, the subjects filled out a questionnaire that included a query of biographical data on items such as age, gender, handedness, the wearing of glasses or contact lenses, amount of experience with HMDs, and the amount of hours spent playing 3D computer games, plus a query to determine the immersive tendency of a person, which was adapted from the presence questionnaire from Scheuchenpflug (2001). The immersive tendency represents a personal trait and describes the readiness of a person to feel immersed in a virtual or remote environment. It can be subdivided into the subscales "emotional involvement" and "degree of involvement," which were both included in the questionnaire. For each of the two subscales, data from a reference sample is available and can hence be compared to the sample at hand (Scheuchenpflug, 2001, 2005).

5.4 Design

There was a one-factorial experimental design with the factor hybrid display (yes versus no), which was investigated between subjects. The dependent variables were the number of identified objects, their estimated distance, the subjects' feeling of telepresence, and the accuracy of the recalled position of the identified objects on the screenshot.

5.5 Results

Detail attention was measured as the number of identified objects from the four objects mentioned above. The difference between the real distance of these objects and the distance estimated by the subjects indicated the goodness of distance estimation. The participants' feeling of telepresence was quantified through an integer between 1 and 7, where a higher value indicates a higher feeling of telepresence. The goodness of recalled position of the identified objects revealed the Euclidean distance between the marked position and the real position of the respective objects on the screenshot.

All dependent variables were averaged across subjects and conditions and analyzed by parametric tests for two independent samples.

The relationship of the dependent variables with the data from the questionnaire was evaluated accessorily in order to control for possible influences other than the experimental variation.

5.5.1 Descriptive Analysis of the Questionnaire Data. None of the biographical data collected showed a relationship to either of the dependent variables.

There was no significant correlation between immersive tendency (in either of the two subscales emotional involvement or degree of involvement) to the quoted feeling of telepresence, which indicates that it was not the personal trait, but rather the experimental variation that led to the rating of the feeling of telepresence in this experiment.

A significant point-biserial correlation between age and wearing of glasses or contact lenses was found, $r_{pb} = .446$, $p = .006$ ($n = 37$). This means that the frequency of wearing glasses or contact lenses permanently or just under special circumstances increases as age increases. It is known that vision is subject to aging processes, so this result is not surprising.

The two subscales of the immersive tendency questionnaire showed a high (Spearman's rank) correlation of $r_S = .716$, $p < .000$ ($n = 37$) among each other.

Based on the fact that the emotional involvement and the degree of involvement are subscales of the same dimension, namely immersive tendency, a high relationship was expected.

Women scored higher on both subscales than men. The scores for women were Md = 31.000 for the emotional involvement and Md = 23.750 for the degree of involvement, in contrast to Md = 23.000 and Md = 18.500 as the corresponding scores for men.

The correlations between gender and emotional involvement (r_{pb} = .460) and between gender and degree of involvement (r_{pb} = .490) turned out to be significant, $p < .004$ (n = 37).

This finding replicates the effect found in the literature that women show higher presence scores in noninteractive virtual environments than men (Nicovich et al., 2005).

It was found that the sample at hand with a mean of M = 24.595 and SD = 6.849 for emotional involvement and M = 19.919, SD = 5.639 for degree of involvement was not different from the reference sample (M = 22.87 for emotional involvement and M = 18.70 for degree of involvement, n = 165; (Scheuchenpflug, 2001; personal communication, Scheuchenpflug & Pongrac, 2005). The confidence interval, where 95% of the sample is located, ranged at the present sample from 22.39 to 26.80 for the emotional involvement and from 18.10 to 21.74 for the degree of involvement. It can be seen that the scores of the reference sample lie within these ranges, which means that the reference sample and the sample at hand do not differ, indicating that the present sample is a representative one.

5.5.2 Number of Identified Objects as Measure for Detail Attention. In the condition with the video view only, the median number of identified objects was Md_V = 1.333 with an average quartile range of Q_V = 1.297 (n = 18). With the additional virtual model displayed, the median and the average quartile range were Md_H = 2.308 and Q_H = 1.674, respectively (n = 19).[1]

1. The median and the average quartile range are reported here because the number of identified objects has more characteristics of an ordinal scale due to the upper limit.

This difference in the number of identified objects between the two conditions turned out to be significant, U = 88.000, p = .009. This means that when the hybrid view is displayed, there is a better detail attention than in the condition with video view, as predicted.

On closer examination, this significant difference was due to identification of the two objects that had been implemented in the virtual model, namely the green box and the No Smoking sign. For these two objects, the average number identified was Md_V = 0.294 and Q_V = 0.824 in the condition with the video view only, versus Md_H = 1.467 and Q_H = 1.313 in the condition with the hybrid view. This advantage of the condition where the virtual model combined with the video view was displayed was significant, U = 63.500, $p < .000$.

However, for the other two objects (the small champagne bottle and the tennis ball), which were not included in the virtual model and could only be seen in the video view, the average number of identified objects in the condition with video view only turned out to be Md_V = 0.867 with deviation Q_V = 1.272, whereas in the condition with the hybrid model the median was Md_H = 0.786 with Q_H = 1.484. This difference was not significant, U = 162.500, p = .781.

There seems to be an effect of peripheral vision in the way that objects, which are displayed continuously in the peripheral field of view, are remembered better.

5.5.3 Estimated Distance. It was predicted that the estimated distance should be better in the condition with the hybrid model displayed than in the condition with video view only. For this reason, the absolute difference between the estimated and the actual distance for the four objects were calculated and averaged across all objects and subjects for the two conditions.

The average difference had a mean of M_V = 0.861 m with a standard deviation of SD_V = 0.516 m in the condition with the video view only, whereas in the condition with the hybrid model the average difference was M_H = 0.452 m, standard deviation SD_H = 0.248 m. The difference between these two conditions turned out to be significant, $t(31)$ = 2.876, p = .009.

Indeed, the subjects were able to estimate the distance more accurately when the video view was combined with the virtual model through the HMD than when it was not. Hence, the presence of the virtual model can mitigate the effect of the diminished distance estimation observed in previous studies having a restricted field of view.

5.5.4 Feeling of Telepresence. The subjects reported a feeling of telepresence in the condition with video view only with an average of $Md_V = 3.778$, with average quartile range $Q_V = 2.064$ ($n = 18$). In the condition with the hybrid model displayed it was $Md_H = 4.800$, $Q_H = 1.875$. This difference between the two conditions in rating of feeling of telepresence was significant, $U = 104.500$, $p = .038$.

This means that the subjects felt more telepresent in the condition with the hybrid model displayed than in the condition with video view only, as predicted.

5.5.5 Accuracy of Recalled Position. This variable was quantified by measuring the Euclidean distance of the marked positions of the identified objects by the participants to the actual positions of these objects on the screenshot. In the condition with video view only, this distance averaged $M_V = 1.995$ cm with a standard deviation $SD_V = 1.764$ cm. In the hybrid model condition, the distance constituted $M_H = 2.760$, $SD_H = 1.994$. This difference turned out not to be significant, $t(31) = -1.165$, $p = .253$. This was contrary to the prediction. Maybe there was an effect of transfer of the subjects' cognitive 3D representation of the scene to the 2D screenshot.

To conclude, the hybrid view led to a higher number of identified objects, to a better distance estimation, and to a higher feeling of telepresence compared to the video view only. This effect can be attributed to the widened field of view.

6 Experiment 2

This experiment was conducted in order to determine which combination composed of the augmentation variant of the virtual model, the transition, and the degree of reprojection error, gives the highest rating of quality, the highest feeling of telepresence, and the best situation awareness. It was of special interest if an optimum, regarding these three dependent variables considered together, could be found.

In this experiment, the camera position was static and the variations of the independent variables—augmentation, blending, and reprojection error—were implemented as stationary, so that the subjects were able to survey the design of the transition between the real and the virtual view.

There were three types of augmentation of the virtual model, namely none, a checker pattern in order to have an abstract version of this variable, and a colored depositation of the virtual view. The color chosen was a light blue (red 173, green 216, blue 230, each on a scale from [0. . .255]) because the virtual view is most often displayed in the peripheral field of view or in the background. It has been shown that light blue is particularly suitable for the peripheral field of view or for background views (Wandmacher, 1993; Schmidtke, 1981).

The transition was introduced in two levels, namely no blending, that is, an abrupt transition from video to virtual view, and blending, that is, a smooth blending only in the border regions of the video view.

The reprojection error results because of translational movements of the human operator due to the lack of depth information from the picture covered by the camera. The levels introduced were 0 m (no reprojection error), 0.05 m, 0.15 m, 0.30 m, and 0.50 m.

The hypotheses regarding the effects of these variables on the rating of quality, the feeling of telepresence and the situation awareness of the subjects were as follows.

- The subjects' rating of quality and feeling of telepresence should be highest with no augmentation of the virtual model; both of these dependent variables should be degraded when the virtual model is augmented with a checker pattern or with the blue color.
- The situation awareness of the subjects should be best when augmentation of the virtual view, either

with the checker pattern or with colored deposita-tion, is presented, and should be worst with no augmentation of the virtual model.

- Blending should enhance the rating of quality and the feeling of telepresence in comparison to an abrupt transition.
- The situation awareness should be increased with an abrupt transition in contrast to blending.
- With incremental reprojection error, the rating of quality and the feeling of telepresence should decline whereas small reprojection errors should be tolerable by the subjects without impairing the rating of quality and the feeling of telepres-ence.
- With increasing reprojection error the situation awareness should increase due to the clear marking of the virtual model with this error.
- In regard to the rating of quality and the feeling of telepresence, the transition and the reprojection error should not develop independently from one another, that is, as reprojection error increases, the effect of transition on these two variables dimin-ishes. This is expected because the reprojection er-ror is a perceptual perspective distortion that out-weighs the transition. This interaction effect should not be true for the situation awareness.
- The reprojection error should serve as a marking of the views just like augmentation. Thus, as reprojec-tion error increases, the impact of augmentation on situation awareness is expected to decrease. This effect should not crop up in the rating of quality and the feeling of telepresence.
- When the reprojection error is small and blending is used, then the augmentation of the virtual view should be necessary to preserve a high degree of situation awareness. This effect should be attenu-ated without blending. With increasing reprojection error, the need for augmentation of the virtual view to preserve a high degree of situation awareness is expected to be higher with blending than without. There should be no such interaction between all three independent variables for the rating of quality and the feeling of telepresence.

6.1 Technical Setup

Similar to Experiment 1, the camera images were sent to the operator, but here the camera position was set to a fixed position. Thereby the operator could focus his or her attention on judging visual quality at the bor-ders of the video view. In order to give a cue as to which part of the display is the video view, a toy train drove in front of the cameras at the remote environ-ment. As this moving train was not part of the 3D rep-resentation it did not drive in the virtual view, but only in the video view, in order to ensure that the subjects were aware of which kind of view they were looking at. The tracked HMD positions were used similarly as they were in Experiment 1 to generate the hybrid view; but to introduce a fixed amount of reprojection error an artificial amount of lateral translation (here to the left side) was introduced between the operator observing position and the camera frame position. In Figures 2, 3, and 4 the levels of the three independent variables are illustrated.

6.2 Subjects

The same subjects used in Experiment 1 partici-pated in Experiment 2.

6.3 Procedure

After carrying out Experiment 1, the subjects pro-ceeded to Experiment 2. Here, the participants looked at a section of the scene shown in Experiment 1, com-posed of a desk with rails on it, for 6 s. This time length was chosen because it has been shown that 4 to 8 s are enough to encompass a 3D visual scene (Lin et al., 2003; Blackmon et al., 1997). This section of the scene combined the video view with the photorealistic virtual model, realizing the different variations regarding aug-mentation of the virtual view, transition, and reprojec-tion error in the stationary way mentioned above. Subjects were instructed to attend to the combination of video and virtual views, whereas the experimenter did not specify a particular dimension so as to avoid

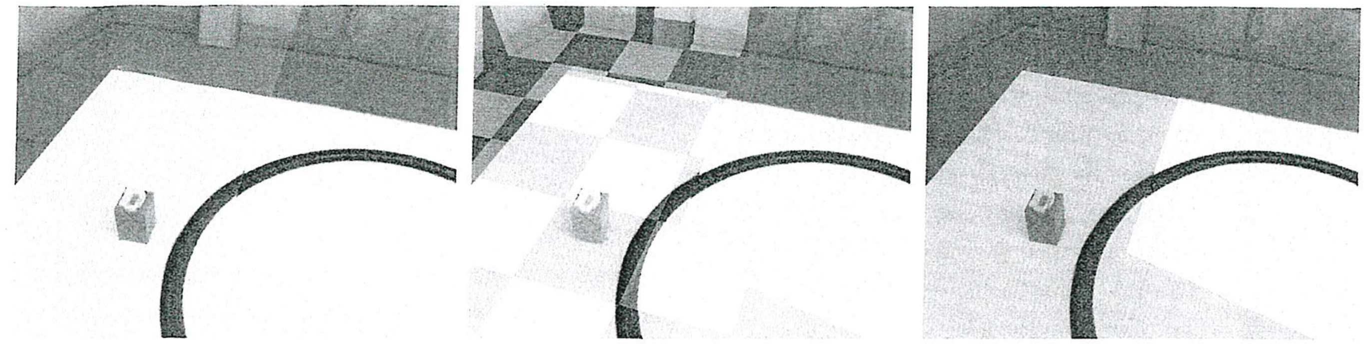


Figure 2. *Levels of view marking: none, checker pattern, and blue color (from left to right).*

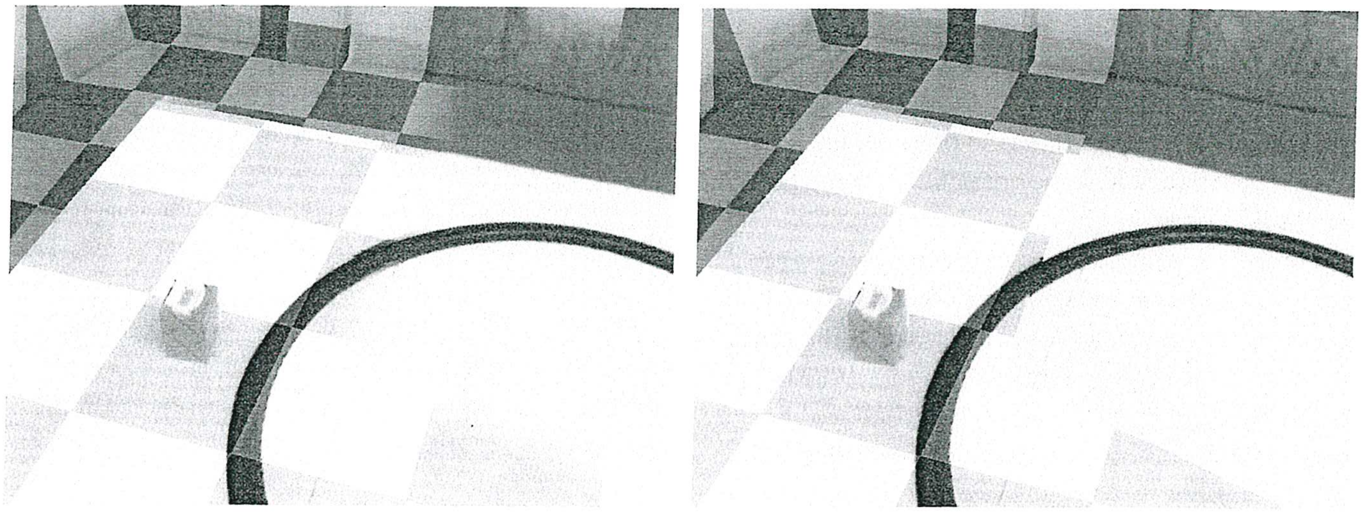


Figure 3. *Levels of transition: blending (left) and no blending (right).*

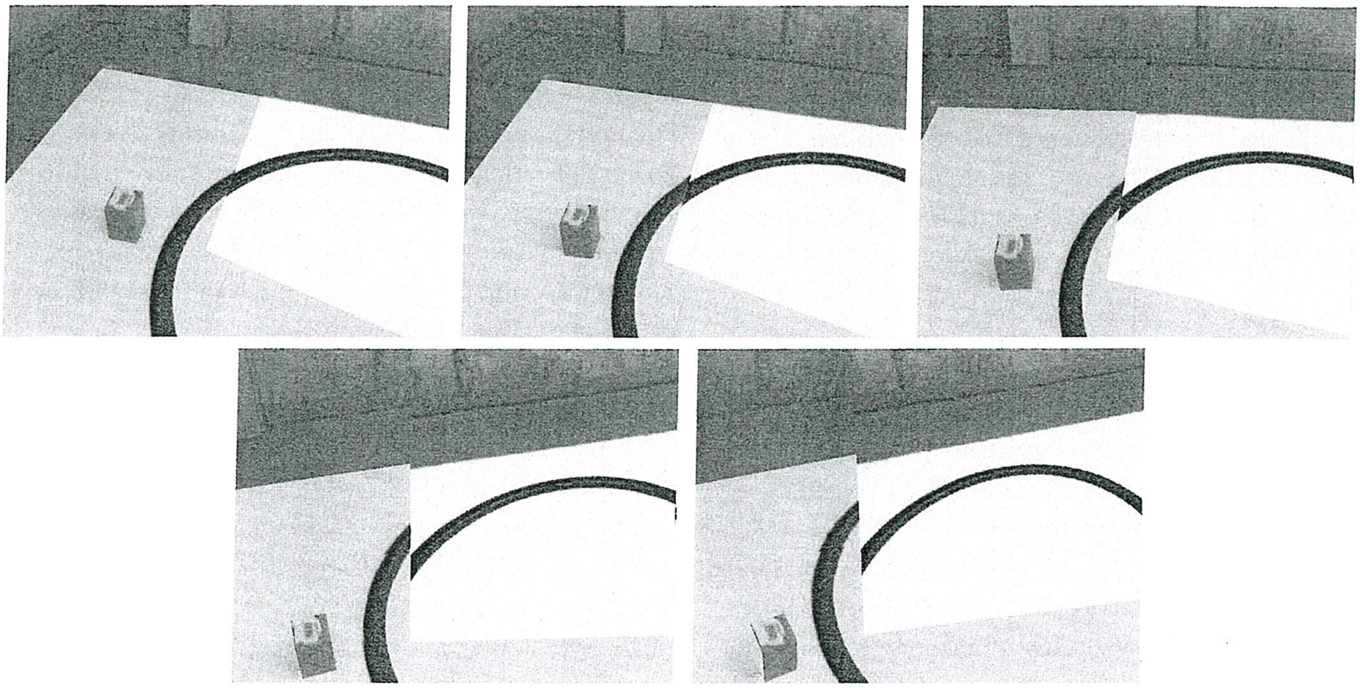


Figure 4. *Levels of reprojection error: 0, 0.05, 0.15, 0.3, and 0.5 m (from left to right, top to bottom).*

directing the subjects' attention to a particular characteristic of the combination.

The first trial was repeated, when necessary, until the subjects knew where to attend. The participants' observation of the views was monitored on a screen.

After each trial, the subjects were asked to rate the quality of the combination and their feeling of telepresence. To assess the rate of quality, the subjects were asked how they assessed the design of the combination on a scale from 1 to 10, where 1 meant very bad and 10 very good. The feeling of telepresence was measured using a question (How strongly did you feel immersed in the shown environment?) with the instruction for scaling from 1 (very weakly) to 7 (very strongly) as in Experiment 1. These two different scale ranges were chosen in order to ensure that the subjects had to rethink their response so that repetition bias was minimized.

In order to obtain the subjects' situation awareness, a beep sounded at, on average, every second trial and the current picture was stopped for 1 s (which is the decay time for the visual image store; Card et al., 1983), then the picture disappeared. Here, every second trial was chosen in order not to overstrain the participants. The subjects' task was to determine the percentage of the video view from the whole view of the picture shown when brought to a stop. The beep sounded at random and at unexpected points in time. After the rating of the percentage of the video view, the same trial was shown again for 6 s in order to get the ratings for the quality and the feeling of telepresence from the subjects.

The sequence of the trials was randomized across the subjects and no consecutive two trials were the same for any subjects.

6.4 Design

There was a full three-factorial design (3 × 2 × 5) with independent variables augmentation (none, checker pattern, and color), transition (no blending vs. blending), and reprojection error (0, 0.05, 0.15, 0.3, and 0.5 m, respectively). All independent variables were investigated within subjects. The dependent variables were rating of quality, feeling of telepresence, and situation awareness.

6.5 Results

As already mentioned, the rating of quality was the value the subjects quoted on a scale from 1 to 10 as well as the feeling of telepresence on a scale from 1 to 7, where a higher value indicates a greater expressiveness of the variables.

Situation awareness was measured by calculating the difference between the percentage of the video view stated by the subjects and the percentage specified by the system for the picture brought to stop. Since the situation awareness has been measured on average every second trial for each person, the missing values are replaced by the respective group means.

The rating of quality, the feeling of telepresence, and the situation awareness were averaged across subjects and across conditions and analyzed by a three-factorial three-variate analysis of variance with repeated measurements.

The relationship of the dependent variables with the data from the questionnaire was evaluated accessorily in order to control for possible influences other than the experimental variation.

6.5.1 Analysis of the Questionnaire Data.

A significant correlation between the wearing of glasses or contact lenses and the rating of quality was found, $r_{pb} = -.440$, $p = .006$, $n = 37$. This indicates that the wearing of glasses or contact lenses permanently or in special circumstances degrades the rating of quality. A one-way analysis of variance (ANOVA) was performed in order to examine whether the three groups (wearing glasses permanently, wearing glasses for special circumstances, and not wearing glasses) are different in respect of the rating of quality. The dependent variable was the rating of quality, averages for each subject. The mean rating of quality wearing no glasses or contact lenses at all was $M = 5.831$, $SD = 1.870$ ($n = 14$). that wearing glasses permanently was $M = 4.639$, $SD = 1.397$ ($n = 18$); and that wearing glasses only in special circumstances was $M = 3.727$, $SD = 1.038$ ($n = 5$). These differences turned out to be significant, $F(2, 34) = 4.117$, $p = .025$. The Bonferroni-test as a post-hoc test for the comparison of the discrete levels of an independent vari-

'able revealed that the level for wearing glasses in special circumstances differed significantly from the other two levels (p = .042), these not being distinct from one another. Thus, the relationship between the wearing of glasses or contact lenses and the rating of quality may be due to the fact that the subjects who wear glasses for special opportunities have a reduced visual acuity and therefore rate the quality worse than subjects who do not wear glasses or have permanently corrected vision. The analysis of the other biographical data, the immersive tendency, and the comparison with the reference sample (Scheuchenpflug, 2001, 2005) have already been reported in Section 5.1.

Concerning the rated feeling of telepresence in Experiment 2, there was no significant correlation with either of the subscales of the immersive tendency (r_S = .107, p = .527, n = 37 between emotional involvement and the feeling of telepresence quoted in Experiment 2; r_S = .086, p = .614, n = 37 between degree of immersion and feeling of telepresence quoted in Experiment 2). Again, the conclusion can be drawn that not the personal trait but the experimental variation leads to the rating of the feeling of telepresence in Experiment 2.

There was a high correlation between the rating of quality and the feeling of telepresence, r_S = .689, $p <$ .000, n = 37. In this experiment there seemed to be a great overlapping of the cognitive processes list that led to the rating of quality and the feeling of telepresence.

6.5.2 General Descriptive Results. The highest rating of quality and the highest feeling of telepresence across all subjects (n = 37) were observed when there was no augmentation of the virtual view, blending, and no reprojection error (0 m), and the respective means were M = 6.973 with a standard deviation of SD = 1.861 for the rating of quality and M = 5.162, SD = 1.259 for the feeling of telepresence. On the other hand, this combination yielded the worst situation awareness with a mean difference between the estimated and the actual percentage of the video view of M = 21.050%, SD = 27.273%, averaged for all subjects.

The best situation awareness was found when the virtual model was augmented with a checker pattern for no blending and 0.5 m reprojection error, the mean difference between estimated and the actual percentage of the video view was M = 6.000%, SD = 5.773% for this combination across all subjects. Here, the rating of quality and the feeling of telepresence across all subjects were worst within all combinations with the respective means and standard deviations of M = 3.486, SD = 2.479 for the rating of quality and M = 2.865, SD = 1.735 for the feeling of telepresence.

6.5.3 Rating of Quality. The overall mean for the rating of quality across all subjects and all conditions was M = 4.967, SD = 1.691.

The descriptive statistics for the dependent variable rating of quality with respect to the different levels of augmentation, transition, and reprojection error are summarized in the Appendix in Table A1.

The F-statistics derived from the multivariate analysis of variance (MANOVA) for the rating of quality are specified in Table 1.

As can be seen from Table A1 in the Appendix and Table 1 here, all predicted main effects are significant in the expected direction, that is, no augmentation of the virtual model, blending, and no or minimal reprojection error led to the highest ratings of quality.

The Bonferroni-test as post-hoc test revealed that the two augmentation types (checker pattern and color) did not differ from one another (p = 1.000) but differed from no augmentation ($p <$ .000). There was no difference between no and 0.05 m reprojection error (p = .208), with these two levels being distinct from all other levels of this variable ($p <$ .005). Thus, a minimal reprojection error of 0.05 m can be tolerated by the subjects without diminishing the rating of quality.

The partial η^2 denoted in Table 1 is a measure for the effect size. Thus it appeared that the reprojection error was the most important factor that influenced the rating of quality.

The interaction effect between transition and reprojection error turned out to be significant as predicted: with no reprojection error the difference between blending and no blending (6.495 vs. 5.468) was higher than with high (0.50 m) reprojection error (4.018 vs. 3.658). Hence, the conclusion can be drawn that the reprojection error indeed overweighs the effects

Table 1. *MANOVA Results for the Rating of Quality*

MANOVA	Sum of squares	$F(df)$	Significance level p	Partial η^2
Augmentation	75.580	18.866 (2,72)*†	<.000	0.344
Transition	152.181	18.150 (1,36)*	<.000	0.335
Reprojection error	697.582	48.466 (4,144)*†	<.000	0.574
Augmentation × transition	4.427	1.659 (2,72)	.197	0.044
Augmentation × reprojection error	10.105	0.871 (8,288)	.542	0.024
Transition × reprojection error	22.977	4.617 (4,144)*	.002	0.114
Augmentation × transition × reprojection error	7.942	0.704 (8,288)†	.625	0.019

*Significant at the $\alpha = 1\%$ level.
†Corrected for violation of the assumed sphericity through Greenhouse-Geisser correction.

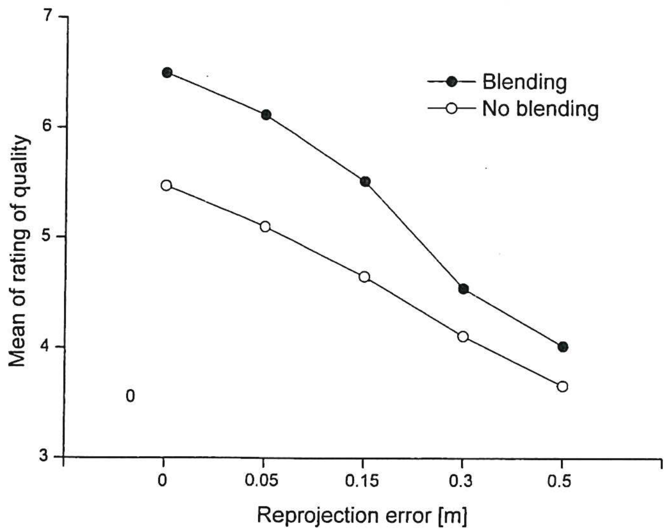


Figure 5. *Interaction effect between transition and reprojection error on rating of quality.*

of the transition on the rating of quality. This interaction effect is illustrated in Figure 5.

No other interaction effect reached significance as expected.

6.5.4 Feeling of Telepresence. The overall mean for the feeling of telepresence the subjects (n = 37) quoted was $M = 3.910$, $SD = 1.179$ across all conditions. The descriptive statistics for the feeling of telepresence in respect to the independent variables are listed in Table A2 in the Appendix.

The F-statistics obtained from the multivariate analysis of variance (MANOVA) for the feeling of telepresence are summarized in Table 2.

Again, all predicted main effects turned out to be significant, the highest feeling of telepresence occurred with no augmentation of the virtual model, blending and no reprojection error.

For the augmentation, the checker pattern and the colored depositation did not differ ($p = 1.000$), whereas they were distinguished from no augmentation ($p <$ .007).

For the reprojection error, all levels differed from one another ($p < .032$), indicating that for the feeling of telepresence not even a small reprojection error was tolerable for the subjects.

Here, the most important variable that influenced the feeling of telepresence was the reprojection error as indicated by the highest effect size for this variable.

In turn, the interaction effect between transition and reprojection error reached significance as for the rating of quality as expected. No reprojection error and blending and no blending yielded means of 4.919 and 4.216, respectively, against high reprojection error (0.50 m) and blending and no blending with means of 3.387 and 3.000, respectively. Similar to the rating of quality, the

Table 2. *MANOVA Results for Feeling of Telepresence*

MANOVA	Sum of squares	$F(df)$	Significance level p	Partial η^2
Augmentation	20.677	7.581 (2,72)*	.001	0.174
Transition	78.933	18.761 (1,36)*	<.000	0.343
Reprojection error	257.000	45.412 (4,144)*†	<.000	0.558
Augmentation × transition	0.180	0.118 (2,72)	.889	0.003
Augmentation × reprojection error	1.989	0.350 (8,288)	.945	0.010
Transition × reprojection error	10.409	4.007 (4,144)*	.004	0.100
Augmentation × transition × reprojection error	4.450	0.994 (8,288)	.441	0.027

*Significant at the $\alpha = 1\%$ level.
†Corrected for violation of the assumed sphericity through Greenhouse-Geisser correction.

reprojection error overweighs the effects of the transition on the feeling of telepresence.

No other interaction effect was significant, as predicted.

6.5.5 Situation Awareness. Situation awareness was operationalized by using the difference between the percentage reported by the subjects when the beep sounded and the picture seen was brought to a stop and the percentage submitted by the system for the same picture. On average, this procedure took place at every second trial. Thus, the higher the value of this difference, the worse the estimation of the subjects was. The overall mean for the situation awareness across all subjects and conditions was $M = 13.406\%$, $SD = 6.435\%$. The minimal deviance averaged for every subject turned out to be 5.40%, and the maximal deviance was 33.80%. Results showed 86.40% of the subjects ($n = 32$) had a divergence of less than 20%, indicating that the subjects were able to satisfactorily estimate the percentage of the video view displayed.

To meet the requirements for MANOVA, the missing values were replaced by the respective group mean for each condition. The following analyses were all conducted with this completed dataset.

The descriptive statistics for the situation awareness in regard to augmentation, transition and reprojection error are shown in Table A3 in The Appendix.

The F-statistics derived from the MANOVA for the situation awareness are listed in Table 3.

As can be seen in Table A3 and Table 3, the difference for the levels of augmentation of the virtual view and transition between the respective group means were significant in the predicted direction.

The Bonferroni-test as a post-hoc test yielded a significant difference between no augmentation and augmentation of the virtual view with the blue color ($p = .001$), that is, the depositation with the color resulted in the best situation awareness. Likewise, no blending lead to the best situation awareness in comparison with blending, as expected. The highest effect size was found for this independent variable, thus the transition was the most important factor influencing situation awareness.

However, although the reprojection error showed a significant effect on situation awareness, it was in the opposite direction to that predicted. Figure 6 makes this clear: There should be a descending line for the means of situation awareness with increasing reprojection error, but the overall tendency is an ascending line, indicating that the situation awareness gets worse with increasing reprojection error. The minimum, that is, the best situation awareness, is located at the reprojection error level of 0.15 m, but the performed Bonferroni-test revealed no significant difference with the reprojection error level of 0 m ($p = 1.000$). These two levels showed a significant difference to the 0.5 m level ($p < .006$),

Table 3. *MANOVA Results for Situation Awareness*

MANOVA	Sum of squares	*F*(*df*)	Significance level *p*	Partial η^2
Augmentation	1200.125	7.894 (2,72)*	.001	0.180
Transition	6671.473	75.216 (1,36)*	<.000	0.676
Reprojection error	2449.137	6.346 (4,144)*†	<.000	0.150
Augmentation × transition	366.705	1.495 (2,72)	.231	0.040
Augmentation × reprojection error	2328.130	3.343 (8,288)*†	.004	0.085
Transition × reprojection error	666.398	1.705 (4,144)†	.170	0.045
Augmentation × transition × reprojection error	5326.232	6.994 (8,288)*†	<.000	0.163

*Significant at the $\alpha = 1\%$ level.
†Corrected for violation of the assumed sphericity through Greenhouse-Geisser correction.

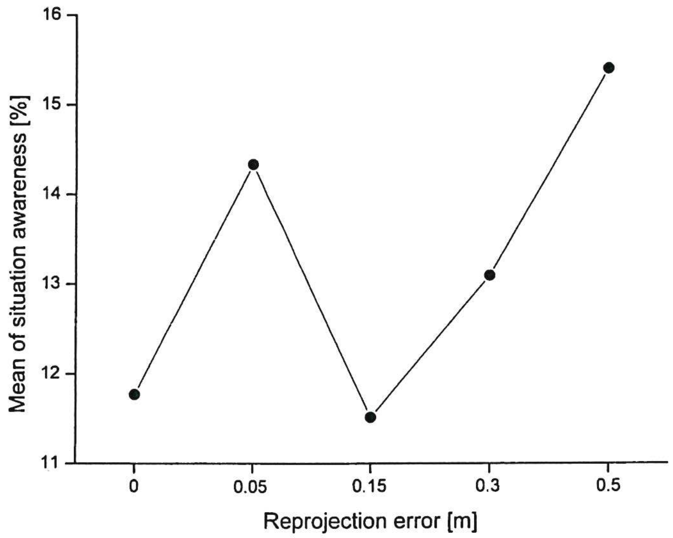


Figure 6. *Main effect of reprojection error on situation awareness.*

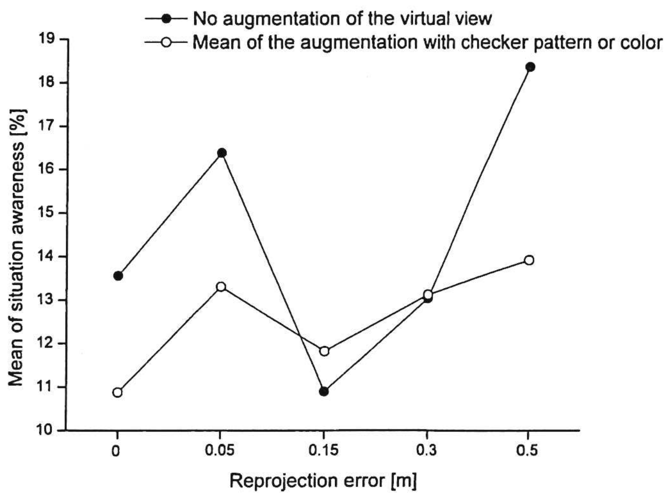


Figure 7. *Interaction effect between augmentation and reprojection error on situation awareness.*

and these were the only significances within the levels of reprojection error regarding situation awareness.

Why this effect emerges is not clear. Perhaps the higher levels of the reprojection error signify such a large perceptual distortion that the subjects were not able to correct the inclined plane for the rating of the percentage of the video view.

It was predicted that the reprojection error serves as a marking of the views just like augmentation with a checker pattern or color, thus making the augmentation of the virtual view unnecessary as reprojection error increases. This should appear as an interaction effect between the reprojection error and augmentation. This interaction took place significantly (see Table 3) and is shown in Figure 7 with the two augmentation types (checker pattern and with blue color) being summarized.

The decreasing difference between no augmentation of the virtual view and augmentation with checker pattern or blue color, respectively, up to a level of 0.3 m of the reprojection error indicates the interaction effect

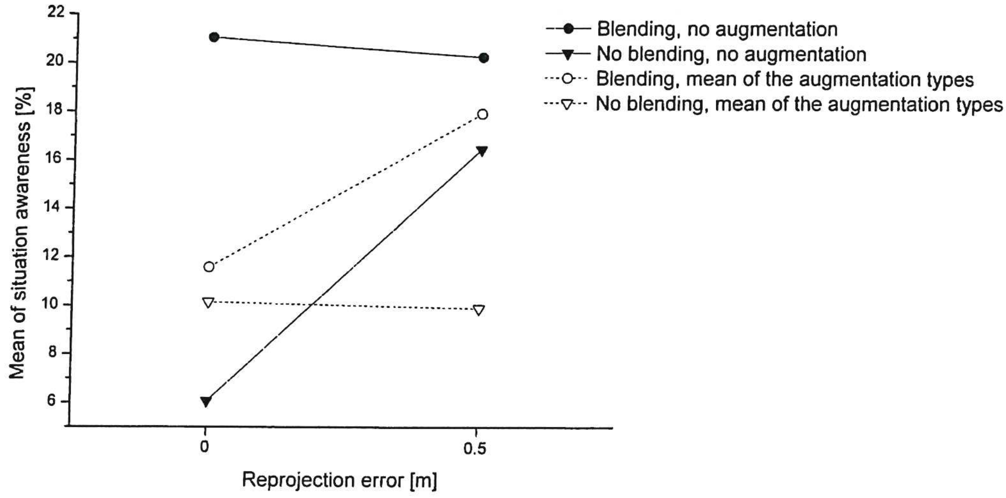


Figure 8. *Three-way-interaction effect between augmentation, transition, and reprojection error on situation awareness.*

that was predicted. Merely the level of 0.5 m of the reprojection error represents an outlier that is not strong enough to suppress the overall interaction effect. Again, there exists the possibility that the subjects were not able to cope with the reprojection error at this high level.

Furthermore, a three-way interaction between augmentation, transition, and reprojection error on situation awareness was expected. This interaction effect turned out to be significant, confirming all predictions made in Section 6 (see Table 3).

Figure 8 illustrates this three-way interaction. For simplification purposes the augmentation types (checker pattern, depositation with the blue color) are summarized again and the reprojection error is reduced to the levels 0 m (none) and 0.5 m because this represents the overall trend.

It was expected that with small reprojection error and blending (lines with the circle symbol in Figure 8), the augmentation of the virtual view with checker pattern or with blue color (dotted lines and open symbols in Figure 8) results in a better situation awareness than when augmentation is left out (compact lines and filled symbols in Figure 8). As can be seen from Figure 8, this was true: the mean difference between the estimated and the video views was $M = 21.050\%$ for 0 m reprojection error, blending, and no augmentation vs. $M = 11.590\%$ for 0 m reprojection error, blending, and augmentation with checker pattern or color.

This effect was expected to be attenuated without blending and this turned out to be the case, too: the mean difference for 0 m reprojection error, no blending, and no augmentation was lower ($M = 6.067\%$) than that with blending; the mean difference for 0 m reprojection error, no blending, and augmentation with checker pattern or blue color was lower ($M = 10.159\%$) than that with blending.

Further, it was predicted that with a high level of reprojection error and blending the augmentation of the virtual view with checker pattern or blue color would, on the one hand, result in a better situation awareness than with no augmentation, but, on the other hand, result in a worse situation awareness than without blending. Again, this effect was confirmed: the mean difference for 0.5 m reprojection error, blending, and augmentation with one of the two types was $M = 17.937\%$ vs. $M = 20.267\%$ for 0.5 m reprojection error, blending, and no augmentation of the virtual view vs. $M = 9.907\%$ for 0.5 m reprojection error, no blending and augmentation of the virtual model.

No other interaction effect turned out to be significant as predicted.

6.5.6 Optimal Combinations of Video and Virtual View. It has been shown that the rating of quality and the feeling of telepresence run opposed to the situation awareness. When the rating of quality and the feeling of telepresence are high, then the situation awareness is bad, and vice versa.

In order to find an optimum between these three variables, the field of application where the hybrid display is applied has to be considered.

When it is important to assess the view as good and to feel highly present, then the combination that yielded the highest rating of quality and the highest feeling of telepresence (no augmentation of the virtual view, blending, and no reprojection error) should be adopted. This could be the case in entertainment applications where little or no manipulation in the remote environment with telepresence systems takes place.

When it is important to have very good situation awareness, such as in manipulation tasks where the handling of tools is required (the tools should not be deposited in the area where the virtual model is displayed when this view is not updated regularly), then the combination that produced the best situation awareness (augmentation of the virtual view with checker pattern, no blending, and a reprojection error of 0.5 m) should be chosen. This can be used for servicing missions in space through telemanipulation where it is extremely important to know what kind of view is being seen, for example.

In order to find optimal combinations of virtual and video view for applications in between these two key points, a median split for the empirical dataset was made for the rating of quality (Md = 4.700) and for the feeling of telepresence (Md = 3.856). All combinations with a higher value were selected.[2] For the situation awareness, all combinations where values of the mean difference were smaller than the first quartile of the sample (25%), which was Q_1 = 11.127%, were selected from the dataset. Treating the whole dataset with this three split procedure, four combinations remained:

When there is a need for a high rating of quality, a high feeling of telepresence, and a sufficient degree of situation awareness, then either the combination with no augmentation of the virtual view, no blending, and no reprojection error or that with augmentation with the blue color, blending, and a level of the reprojection error of 0.05 m, should be selected. Here, the means of the situation awareness were 10.429% for the former combination and 10.500% for the latter combination. The means for the rating of quality were above 5.8, and for the feeling of telepresence above 4.6.

With the demand for a good situation awareness and a sufficient degree of rating of quality and feeling of telepresence, then either the combination with augmentation with checker pattern, blending, and no reprojection error or that with augmentation with depositation of the virtual view with the blue color, no blending, and no reprojection error, should be used. The former showed a mean for the situation awareness of 6.067%, the latter of 8.000%. The means for the rating of quality was above 5.2, and for the feeling of telepresence above 4.0.

7 Conclusions/Discussion

In this article it has been shown that the combination of live video streams and virtual views can improve human performance in telepresence scenarios. A telepresence system was evaluated that widens the available camera field of view by using virtual reality for the peripheral regions that are not covered by the cameras.

In the first experiment this widened field of view lead to better attention to detail, a better distance estimation as well as an enhanced feeling of telepresence of the subjects.

However, the accuracy of the recalled position of objects was not ameliorated as expected. There exists the possibility that this was due to shortcomings of the operationalization of this variable in this experiment.

The conclusion can be drawn that the integration of virtual reality with video data is a viable solution to generate wide fields of view even with restricted camera op-

2. A median and quartile split, respectively, are common statistic procedures for finding the optimum for diverse variables with contradictory values.

tics in order to increase the human performance as well as the feeling of telepresence.

In the second experiment, the three parameters of view marking, transition (blending/no blending), and reprojection error were selected as criteria for the combination of live video streams and virtual views.

Here, it was examined how these three parameters affect the rating of quality, the feeling of telepresence, and the situation awareness of the participants. It was expected that a combination that encompasses high realism would lead to the highest rating of quality and to the highest feeling of telepresence, but to the worst situation awareness. This prediction was confirmed.

On the other hand, combinations that suffer from realism led to the best situation awareness, but to the lowest rating of quality and the lowest feeling of telepresence as was predicted.

Reprojection error was the sole exception regarding situation awareness. Instead of getting better with increasing reprojection error, the situation awareness was best with small values of this error. Maybe the reprojection error was an effect of the perceptual distortion, which would lead the participants to not be able to compensate for the incline of the displayed view. This result should be reconsidered in further studies with an alternative operationalization of the situation awareness.

In applications where situation awareness is important, such as in telemanipulation tasks, view marking can effectively provide a strong cue. It can be implemented with augmentation techniques in many facets, with two examples evaluated here. Other types of view marking could be investigated in further studies.

In regard to several application fields, optimal transition strategies concerning the trade-off between situation awareness and display quality were discussed. The focus should be on the respective importance of the situation awareness, the quality rating, and the feeling of telepresence adherent in the application. These optima were deduced from the results found in this experiment without direct evaluation. This could be done in further studies, where the found optimal transitions could be proved against each other using different application fields.

In this article, only the case without a moving human operator was considered. Further research aims at expanding and comparing the reported results to telepresence scenarios where the human operator has to move and to navigate through the remote environment.

Acknowledgments

This work was supported in part by the German Research Foundation (DFG) within the Collaborative Research Center SFB 453 on "High-Fidelity Telepresence and Teleaction."

References

Alfano, P., & Michel, G. (1990). Restricting the field of view: Perceptual and performance effects. *Perceptual and Motor Skills, 70,* 35–45.

Azuma, R. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments, 6*(4), 335–385.

Azuma, R., & Furmanski, C. (2003). Evaluating label placement for augmented reality view management. *Proceedings of the 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '03),* 66–75.

Bell, B., & Feiner, S. (2000). Dynamic space management for user interfaces. *Proceedings of the UIST '00: ACM Symposium on User Interface Software and Technology,* 239–248.

Bell, B., Feiner, S., & Hllerer, T. (2002). Information at a glance. *IEEE Computer Graphics and Applications, 22*(4), 6–9.

Blackmon, T., Ho, Y., Matsunaga, K., Yanagida, T., & Stark, L. (1997). Eye movements while viewing dynamic and static stimuli. *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 6,* 2814–2819.

Burkert, T., Leupold, J., & Passig, G. (2004). A photorealistic predictive display. *Presence: Teleoperators and Virtual Environments, 13*(1), 22–43.

Card, S., Moran, T., & Newell, A. (1983). *The psychology of human-computer interaction.* Mahwah, NJ: Lawrence Erlbaum.

Chios, P., & Linney, A. (2004). The design process of an autostereoscopic viewing interface for computer-assisted microsurgery. *IEEE International Conference on Systems, Man and Cybernetics, 1,* 136–143.

Counsell, J. (2004). Interacting with real objects in virtual space. *Proceedings of the 8th International Conference on Information Visualisation (IV '04),* 821–826.

Creem-Regehr, S., Willemsen, P., Gooch, A., & Thompson, W. (2003). The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual environments. Technical report UUCS-03-016, School of Computing, University of Utah. Available at http://www.cs.utah.edu/research/techreports.shtml. Last accessed 16 March 2006.

Dai, F. (2001). Remote supervision and task-level control of industrial robot systems using augmented vision. *Automatisierungtechnik, 49*(7), 329–335.

De Bruyn, B. (1997). Blending transparent motion patterns in peripheral vision. *Vision Research,* 37(5), 645–648.

Ellis, S., & Menges, B. (1998). Localization of virtual objects in the near visual field. *Human Factors,* 40(3), 415–431.

Endsley, M. (2000). Theoretical underpinnings of situation awareness: A critical review. In Endsley, M. R. and Garland, D. J. (Eds.), *Situation Awareness. Analysis and Measurement* (pages 3–32). Mahwah, NJ: Lawrence Erlbaum.

Furmanski, C., Azuma, R., & Daily, M. (2002). Augmented-reality visualizations guided by cognition: Perceptual heuristics for combining visible and abscured information. *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR '02),* 215–224.

Hua, H., Brown, L., & Gao, C. (2004). SCAPE: Supporting stereoscopic collaboration in augmented and projective environments. *IEEE Computer Graphics and Applications, 24*(1), 66–75.

Ilie, A., Low, K.-L., Welch, G., Lastra, A., Fuchs, H., & Cairns, B. (2004). Combining head-mounted and projector-based displays for surgical training. *Presence: Teleoperators and Virtual Environments, 13*(2), 128–145.

Insko, B. (2003). Measuring presence: Subjective, behavioral, and physiological methods. In G. Riva, F. Davide, and W. IJsselsteijn (Eds.), *Being there: Concepts, effects and measurement of user presence in synthetic environments,* chapter 7. Amsterdam: IOS Press.

Jay, C., & Hubbold, R. (2003). Amplifying head movements with head-mounted displays. *Presence: Teleoperators and Virtual Environments, 12*(3), 268–276.

Leupold, J., & Behrendt, S. (2005). The hybrid display: Smooth transition between virtual and real views for location unbound observations in telepresence applications. *Joint International COE/HAM-SFB453 Workshop on Human Adaptive Mechatronics and High Fidelity Telepresence,* 25–30.

Lin, S.-K., Seibel, E., & Ill, T. (2003). Testing visual search performance using retinal light scanning as a future wearable low vision aid. *International Journal of Human-Computer Interaction, 15*(2), 245–263.

Livingston, M. (2005). Evaluating human factors in augmented reality systems. *IEEE Computer Graphics and Applications, 25*(6), 6–9.

Low, K.-L., Ilie, A., Welch, G., & Lastra, A. (2003). Combining head-mounted and projector-based displays for surgical training. *Proceedings of the IEEE Virtual Reality (VR '03),* 110–117.

Low, K.-L., Welch, G., Lastra, A., & Fuchs, H. (2001). Life-sized projector-based dioramas. *Proceedings of the ACM Symposium on Virtual Reality and Software Technology 2001,* 93–101.

Macchiarella, N., & Vincenzi, D. (2004). Augmented reality in a learning paradigm for flight and aerospace maintenance training. *The 23rd Digital Avionics Systems Conference, 1*(24–28), 5.D.1-1–5.D.1-9.

Nakatsuru, T., Yokokohji, Y., Eto, D., & Yoshikawa, T. (2003). Image overlay on optical see-through displays for vehicle navigation. *Proceedings of the 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '03),* 286–287.

Nash, E., Edwards, G., Thompson, J., & Barfield, W. (2000). A review of presence and performance in virtual environments. *International Journal of Human-Computer Interaction, 12*(1), 1–41.

Nicovich, S. G., Boller, G. W., & Cornwell, T. B. (2005). Experienced presence within computer-mediated communications: Initial explorations on the effects of gender with respect to empathy and immersion. *Journal of Computer-Mediated Communication, 10*(2), article 6. Available at http://icmc.indiana.edu/vol10/issue2/nicovich.html. Accessed 19 June 2007.

Ou, S., Karuppiah, D., Fagg, A., & Riseman, E. (2004). An augmented virtual reality interface for assistive monitoring of smart spaces. *Proceedings of the 2nd IEEE Annual Conference on Pervasive Computing and Communications (PERCOM '04),* 33–42.

Paloc, C., Carrasco, E., Macia, I., Gomez, R., Barandiaran, I., Jimenez, J., et al. (2004). Computer-aided surgery based on auto-stereoscopic augmented realiy. *Proceedings of the 8th International Conference on Information Visualisation (IV '04),* 189–193.

Psotka, J., Lewis, S., & King, D. (1998). Effects of field of view on judgments of self-location: Distortions in distance estimations even when the image geometry exactly fits the

field of view. *Presence: Teleoperators and Virtual Environments, 7*(4), 352–369.
Raskar, R., Brown, M., Yang, R., Chen, W.-C., Welch, G., Towles, H., et al. (1999). Multi-projector displays using camera-based registration. *Proceedings of the IEEE Conference on Visualisation 1999,* 161–168.
Reitsma, P., & Pollard, N. (2003). Perceptual metrics for character animation: Sensitivity to errors in ballistic motion. *ACM Transactions on Graphics, 22*(3), 537–542.
Rosenblum, L., & Macedonia, M. (2005). Tangible photorealistic virtual museum. *IEEE Computer Graphics and Applications, 25*(1), 15–17.
Scheuchenpflug, R. (2001). Measuring presence in virtual environments. In M. Smith, G. Salvendy, and M. Kasdorf, (Eds.), *HCI International 2001* (pages 56–58).
Scheuchenpflug, R. (2005) Personal communication. University of Wuerzburg, Germany.
Schmidtke, H. (1981). *Lehrbuch der Ergonomie (2. Aufl.)* (pages 410–415). Munich: Carl Hanser Verlag.
Semmlow, J., Gauthier, G., & Vercher, J.-L. (1990). Identification of peripheral visual images in a laterally restricted gaze field. *Perceptual and Motor Skills, 70,* 175–194.
Tang, W.-K., Wong, T.-T., & Heng, P.-A. (2005). A system for real-time panorama generation and display in tele-immersive applications. *IEEE Transactions on Multimedia, 7*(2), 280–292.
Thompson, R., Reid, I., Munoz, L., & Murray, D. (2001). Providing synthetic views for teleoperation using visual pose tracking in multiple cameras. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 31*(1), 43–54.
Waller, D. (1999). Factors affecting the perception of interobject distances in virtual environments. *Presence: Teleoperators and Virtual Environments, 8*(6), 657–670.
Wandmacher, J. (1993). *Software-Ergonomie* (pages 314–321). Berlin: de Gruyter.
Willemsen, P., & Gooch, A. (2002). An experimental comparison of perceived egocentric distance in real, image-based, and traditional virtual environments using direct walking tasks. Technical Report UUCS-02-009, School of Computing, University of Utah. Available at http://www.cs.utah.edu/research/techreports.shtml. Last accessed 16 March 2006.
Witmer, B., & Kline, P. (1998). Judging perceived and traversed distance in virtual environments. *Presence: Teleoperators and Virtual Environments, 7*(2), 144–167.

Appendix

Table A1. *Descriptive Statistics for the Rating of Quality*

Independent variable	Level	*M*	*SD*
Augmentation	None	5.335	1.728
	Checker pattern	4.765	1.734
	Blue color	4.800	1.734
Transition	Blending	5.337	1.728
	No blending	4.596	1.813
Reprojection error	None (0 m)	5.982	1.484
	0.05 m	5.608	1.697
	0.15 m	5.081	1.807
	0.30 m	4.324	1.953
	0.50 m	3.838	2.135

Table A2. *Descriptive Statistics for the Feeling of Telepresence*

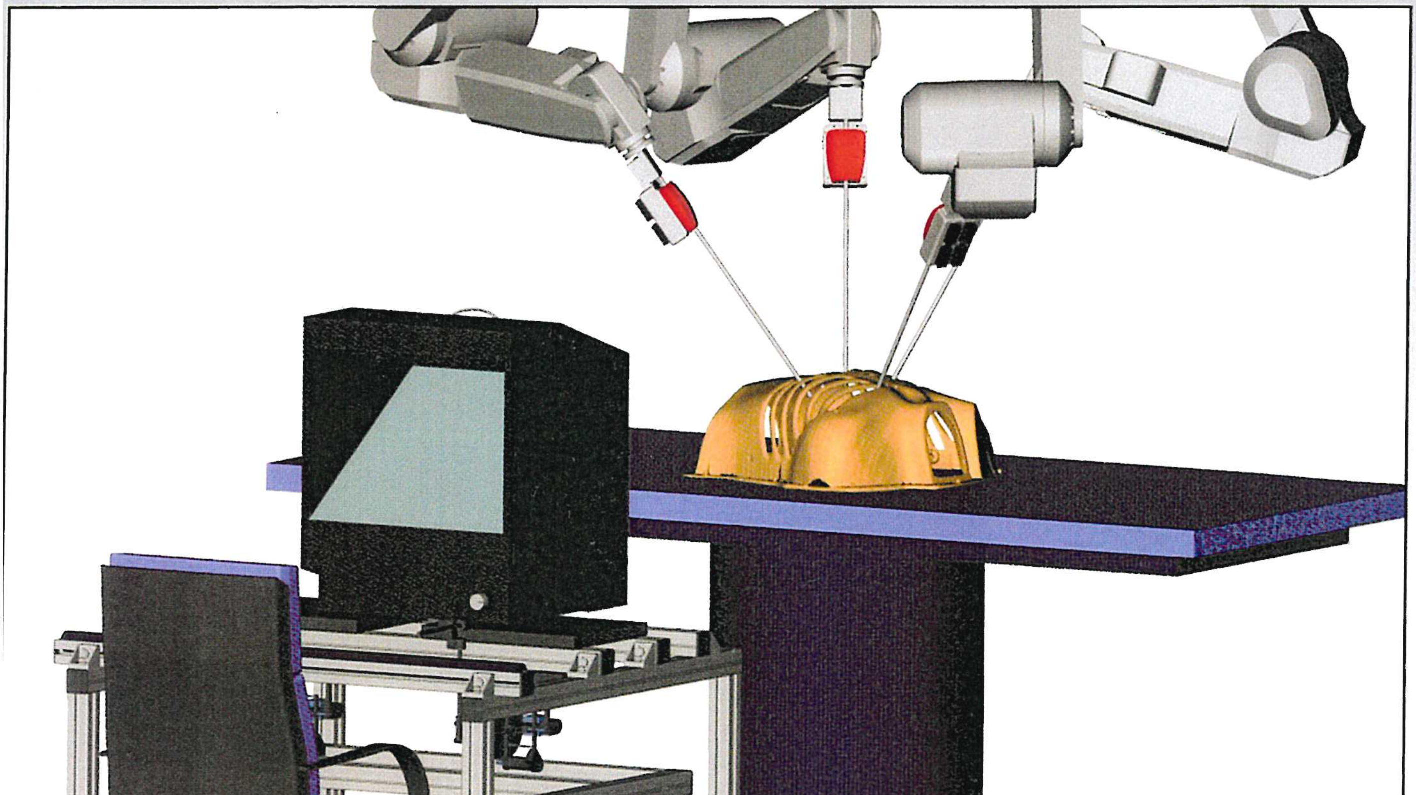
Independent variable	Level	*M*	*SD*
Augmentation	None	4.103	1.168
	Checker pattern	3.822	1.235
	Blue color	3.805	1.247
Transition	Blending	4.177	1.204
	No blending	3.643	1.265
Reprojection error	None (0 m)	4.568	1.089
	0.05 m	4.243	1.162
	0.15 m	3.950	1.265
	0.30 m	3.595	1.387
	0.50 m	3.194	1.350

Table A3. *Descriptive Statistics for the Situation Awareness*

Independent variable	Level	*M* [%]	*SD* [%]
Augmentation	None	14.450	4.252
	Checker pattern	13.311	3.759
	Blue color	11.907	3.645
Transition	Blending	15.674	3.522
	No blending	10.771	3.698
Reprojection error	None (0 m)	11.769	4.526
	0.05 m	14.334	5.256
	0.15 m	11.515	4.440
	0.30 m	13.094	3.990
	0.50 m	15.402	5.566

PRESENCE

TELEOPERATORS AND VIRTUAL ENVIRONMENTS



VOLUME 16 / NUMBER 5 / OCTOBER 2007

MIT PRESS