REGULAR CONTRIBUTED PAPER

WILEY

Light-field head-mounted displays reduce the visual effort: A user study

Michael Panzirsch [ORCID] | Bernhard Weber | Nicolai Bechtel | Nicole Grabner | Martin Lingenauber

Institute of Robotics and Mechatronics, German Aerospace Center (DLR), Wessling, Germany

Correspondence
Michael Panzirsch, Institute of Robotics and Mechatronics, German Aerospace Center (DLR), Wessling, Germany.
Email: michael.panzirsch@dlr.de

Funding information
European Union's Horizon 2020 research and innovation program, Grant/Award Number: 951989

Abstract

Head-mounted displays (HMDs) allow the visualization of virtual content and the change of view perspectives in a virtual reality (VR). Besides entertainment purposes, such displays also find application in augmented reality, VR training, or tele-robotic systems. The quality of visual feedback plays a key role for the interaction performance in such setups. In the last years, high-end computers and displays led to the reduction of simulator sickness regarding nausea symptoms, while new visualization technologies are required to further reduce oculomotor and disorientation symptoms. The so-called vergence–accommodation conflict (VAC) in standard stereoscopic displays prevents intense use of 3D displays, so far. The VAC describes the visual mismatch between the projected stereoscopic 3D image and the optical distance to the HMD screen. This conflict can be solved by using displays with correct focal distance. The light-field HMD of this study provides a close-to-continuous depth and high image resolution enabling a highly natural visualization. This paper presents the first user-study on the visual comfort of light-field displays with a close-to-market HMD based on complex interaction tasks. The results provide first evidence that the light-field technology brings clear benefits to the user in terms of physical use comfort, workload, and depth matching performance.

KEYWORDS

light-field displays, vergence–accommodation conflict, virtual reality

1 | INTRODUCTION

Although the principles of eye vergence[1,2] and the interrelations of stereo vision and depth perception[3] were observed and studied already a long time ago, the physically correct or natural visualization of virtual scenes remains a challenge in modern display technology.

Stereoscopic head-mounted displays (HMDs) provide two views—one for each eye—from view perspectives with a distance corresponding to that of the human eyes. Head-mounted near-eye displays are used for entertainment purposes as movies or gaming and also for industrial virtual reality (VR) training or tele-operation systems. In these applications, for instance, a human

operator uses an input device to control a remote or virtual robotic system to operate in the remote or virtual environment respectively. This requires sufficient image resolution and framerate and also a natural visualization to allow for lengthy usage without visual discomfort.

In the last decade, stereoscopic displays were optimized in terms of resolution, tracking accuracy, and visualization latency with focus on reducing simulator sickness. Thereby, different theories of simulator sickness were considered,[4] which are mainly related to nausea and disorientation symptoms.[5] Besides the remaining simulator sickness issues related to the naturalness of motion cues from input devices,[4] the optimization of standard stereoscopic displays reaches its limits. One remaining major drawback of conventional HMDs is their inability of presenting the content in a physically natural way since all objects are presented with a fixed focal distance. Hence, the content is only visualized sharply when the eyes' lenses focus this distance and, as a consequence, also no variable depth sharpness can be presented. This limitation leads to the so-called vergence–accommodation conflict (VAC,[6-8]). The vergence distance is the distance in which the eyes' lines of sight intersect, while the accommodation distance describes the distance which the eye lens must focus to produce a sharp retinal image.[9] The VAC origins from the mismatch between the 3D pose of an object (vergence distance) and the eye focusing distance (accommodation distance) in conventional stereoscopic displays. The neural coupling of the vergence and accommodation and the respective conflict in conventional stereoscopic displays lead to distortions in 3D perception, difficulty in focusing, and visual discomfort or fatigue.[9] To this end, the VAC is mainly related to the disorientation and oculomotor symptom cluster of Kennedy et al.[5] The focal distance of conventional stereoscopic displays is mostly around 1.5 to 2 m. Thus, the VAC is very pronounced for objects in distances below approximately 1 m regarding the analysis of Banks et al.[10] This renders conventional displays uncomfortable especially for typical human arm reaching distances and thus for interaction tasks.

The light-field or varifocal display technology in general aims to cure this VAC through physically natural presentation of the content. In contrast to conventional stereoscopic displays with one fixed focal distance, light-field displays are able to render different focal planes visible to the human eye. Thus, light-field displays allow focusing objects at the correct 3D position and consequently the VAC can be solved. The authors of Banks et al.[11] reported that predator animals have special pupil shapes for increased depth sharpness which improves the depth perception of the animals. This indicates that light-field displays promise not only to solve the VAC but also to improve the depth perception due to increased depth sharpness.

A large variety of technologies to display different depth areas[12-14] was proposed, including holographic displays[15] and such based on microlens arrays, multiple depth planes,[16,17] or varifocal elements (e.g., focus-tunable lenses[18,19] or deformable membrane mirrors[20]). It has to be noted that most of such displays providing high resolution are still restricted to non-wearable near-eye displays. Other methods apply eye-tracking for eye gaze recognition and the corresponding refocusing of the displayed view.[21] In general, hardware-based solutions as displays that present a full light-field are physically more natural in that the refocusing effect happens in an optically realistic manner and thus instantaneously. In contrast, software-based solutions based on eye-tracking are potentially affected by latency or inaccuracies since sensor and processing loops are involved. While the varifocal devices of the companies Magic Leap (two depth planes,[22]), Microsoft (one depth plane,[23]), or nreal[24] achieve partially sufficient image resolution, the image depth is restricted to only few depth planes as for the devices of vuzix (one depth plane,[25]), Google (one depth plane,[23]), AVEGANT (three depth planes,[26]) and LightSpace Technologies (four depth planes,[27]) with reduced image resolution. On the other side, devices providing up to infinite depth resolution often provide low image resolution (VIVIDQ,[28] CYVISION[29] and SeeReal Technologies[30]).

The light-field head-mounted display[31] applied in this study is not based on microlens arrays, but on sequential projection of different view points as described later in more detail. This solution provides a high depth and spatial resolution at the same time and avoids high computational efforts, eye-tracking, and computation delays. In the presented study, for the first time, the effect of light-field HMDs on the visual comfort and depth matching performance could be evaluated in an application-related setup.

In literature, the presentation of depth information was evaluated with a variety of study concepts to evaluate hardware displaying different depth areas. Still, the effect of light-field rendering and the resulting avoidance of the VAC on the visual comfort was not investigated in detail, so far. In Battisti et al.,[32] pseudo-videos with different sub-aperture view trajectories were evaluated purely based on visual analysis. The pseudo-video was generated from pictures along different trajectories through the sub-aperture views. In Dunn et al.,[20] a monocular user study was performed with a binocular light-field display based on deformable membrane mirrors such that the VAC could not be studied. The user study task was based on solving mathematical equations without complex

interactions. An interactive user-study was performed on a flexible light-field smartphone with microlens array in Gotsch et al.[33] The participants had to position a three-dimensional (3D) object in three dimensions; however, the study investigated the effects of different input modalities, but not the effects of the light-field technology itself. Also, different light-field rendering effects on user acceptance were evaluated in look-and-feel studies.[34]

The presented study evaluates the effects of the light-field head-mounted display technology on the visual effort and comfort respectively. Specifically, we compared conditions with light-field activated (*LF on*) with light-field deactivated (*LF off*). The study was implemented in a virtual reality setup with realistic 3D pick and place as well as abstract 1D depth matching tasks using a haptic input device. Besides objective performance data, subjective data was collected to evaluate the effects of light-field displays on simulator sickness, workload and the system usability.

The following hypotheses were formulated:

H1 Light-field technology eliminates the VAC and consequently reduces the visual effort of the user.

The light-field technology allows for the perception of different depth layers and objects are exposed more clearly in front of more distant objects. Therefore, we hypothesize the following:

H2 The more natural depth presentation increases the user's confidence in depth estimation and reduces the required concentration as well as workload.

No formal hypothesis is formulated wrt to task performance (required time and accuracy). Although a general beneficial effect on accuracy in depth positioning is expected, this effect might be confounded, for example, due to limited depth resolution of the device.

The paper is structured as follows: Section 2 introduces the materials and methods including the description of the experimental setup, study procedure, and metrics. The results are presented in Section 3 and discussed together with the limitations of the study in Section 4. Finally, Section 5 concludes the work.

2 | MATERIALS AND METHODS

2.1 | Sample

$N = 20$ (9 females, 11 males) subjects with an average age of $M = 28$ years ($SD = 4.6$ years; range: 21–38 years)

participated in the study. Ten participants required visual acuity correction and thus wore their glasses or contact lenses during the experiment. None of the subjects indicated to have problems with depth perception.

2.2 | Apparatus

2.2.1 | Hardware

Head-mounted display

Figure 1 depicts the light-field HMD Zorya of the company Creal. The display can be strapped to the head and adjusted with the knob at the back of the device (depicted in Figure 2) and the top strap (hook and loop fastener shown in Figure 3). Figure 4 presents a knob on top of the device with which the inter-pupillary distance (IPD), that is, the individual distance between the two eyes, can be adapted. The lens system (see Figure 5) is suitable to achieve an IPD between 58 and 72 mm.

Figure 6 presents a look-through image of the Zorya HMD as visible to one eye of the user. The peripheral background display is a standard LCD screen with a
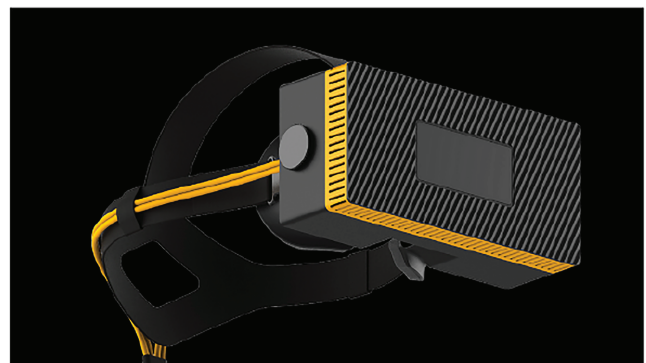

FIGURE 1 Zorya head-mounted display (HMD) (Creal)
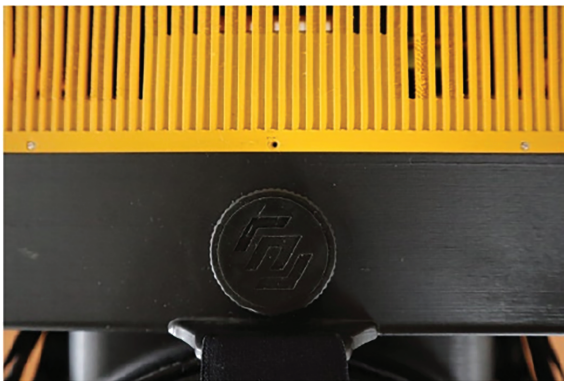

FIGURE 2 Strap knob

FIGURE 3   Top strap

FIGURE 4   Knob for inter-pupillary distance adjustment

FIGURE 5   Lens system

FIGURE 6   Look-through photo

resolution of 1600 × 1440 px and a field-of-view (FOV) of approximately 100° diagonally. It provides the peripheral part of a human's FOV. The central light-field display has a FOV of 30° with approximately 40 px/° angular resolution. The light-field visualization is achieved through an array of 32 pinholes resulting in a continuous depth perception. Due to the small eye box (approximately 8 mm) of the light-field display, it is of high importance that the display is well centered horizontally and vertically with respect to the head. Still, since a wrong offset would be clearly visible, this display configuration helps novel users to position the device on the head correctly for optimal LF visualization. In CREAL's system, the always-in-focus images are created by a sequential pin-light illumination of a fast, spatial light modulator that reflects modulated light beams to imaging optics and toward specific viewpoints in the vicinity of the eye pupil. The modulator is technically a selective mirror that casts a
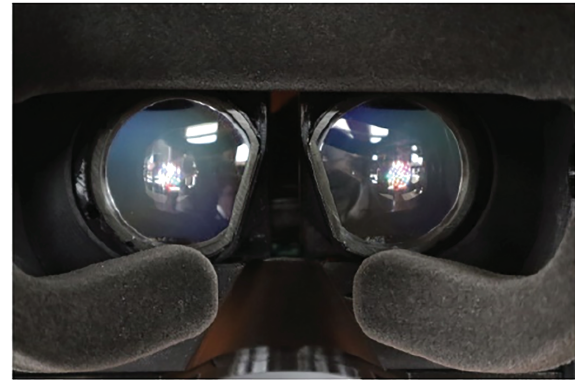
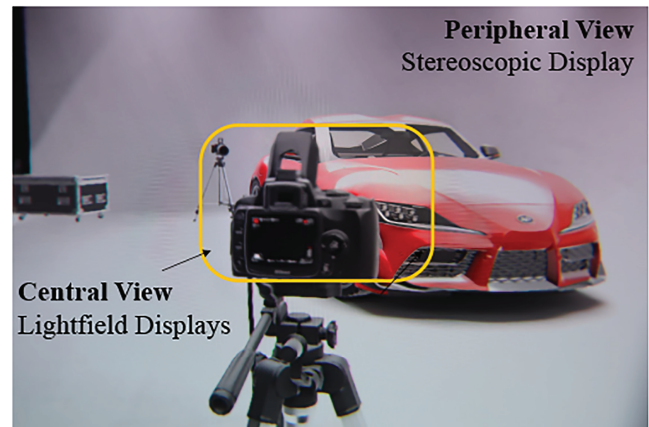shadow of the scene as it is supposed to be seen from the perspective of the particular viewpoint. Figure 7 presents this technical setup, the light beam path for two different viewpoints, and the resulting optical impression for the user. The small size of the pin-light-source causes that the image has a large depth-of-field, that is, the image is practically always-in-focus, and passes seemingly through a virtual pinhole hanging in the air near the eye pupil. Different pin-lights perform the same operation in sequence, but the modulator reflects images of slightly different perspectives of the 3D scene and projects each through a different viewpoint.

A fast sequence of always-in-focus images passing through a 2D array of viewpoints represents a light-field that entirely or almost entirely enters the eye pupil (as visualized in Figure 8). The eye can then focus on virtual objects in any distance. This operation is performed purely by the eye, no eye-tracking is needed. Figure 8 shows that depending on the focus of the eye, the objects in other depth planes appear blurry to the user.

Since the peripheral background display provides a clearly lower resolution than the central light-field
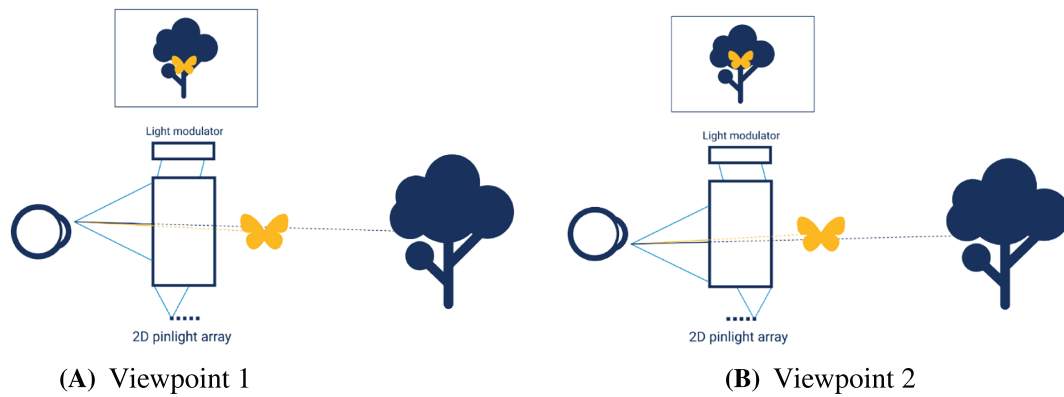
(A) Viewpoint 1   **(B)** Viewpoint 2

FIGURE 7   Light-beam path and user view

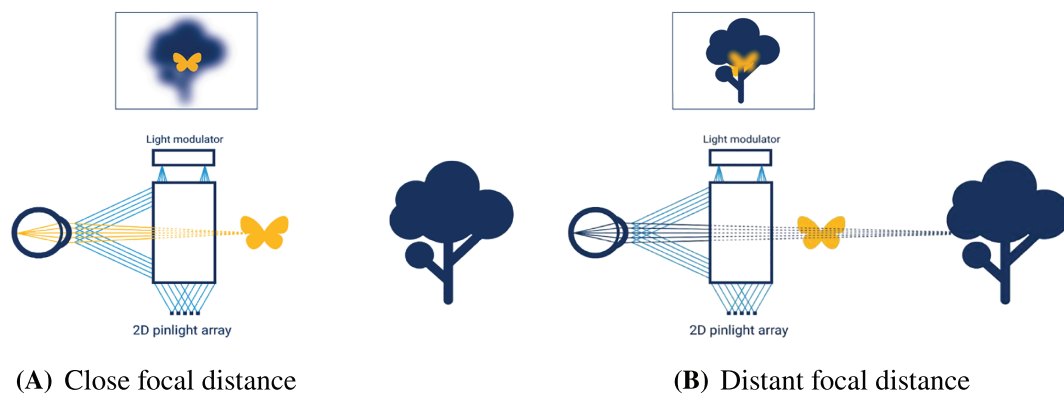**(A)** Close focal distance   **(B)** Distant focal distance

FIGURE 8   Creal's light-field technology

displays, it could not be used as the *LF off* condition. Via software, the light-field display can present either a light-field (*LF on*) or be flattened at an arbitrary focal distance (*LF off*). This flattened display corresponds to a standard stereoscopic display with fixed focal distance. Thus, for both conditions of the study, the high-resolution central light-field displays could be applied in combination with the peripheral background display.

Input device

In order to allow the user to interact in the VR scenes, the haptic interface lambda.7 by Force Dimension (see Figure 9) was used. Figure 10 presents its integration in the experimental setup which is explained later in more detail. The device provides a large translational workspace (ø240 mm × 170 mm), which was further extended through a scaling of 10 in order to allow working in the range of 1.15 m in the VR environment comfortably. The device haptically guided the user to an initial position after

having activated the gripper. From this position the required workspace of each task could be reached.

2.2.2 | Software

The virtual environment and the rendering for the HMD was implemented with the VR-Engine *Unity* and Creal's STALF rendering software. The VR was executed on the Windows host computer of the HMD. The control software was implemented in Matlab/Simulink and executed on a Linux computer at 1 kHz.

2.3 | Experimental setup

Figure 10 presents the experimental setup with the HMD, the input device and a foot pedal. Participants were asked to rest their elbows on the tabletop to avoid too much physical effort during the experiment. Holding the HMD served reducing the physical load on the head
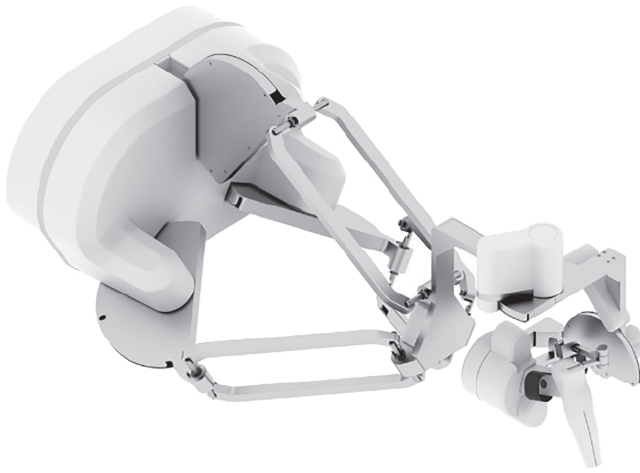
Head Mounted Display
Zorya (Creal)

Haptic Interface
lambda.7 (Force Dimension)

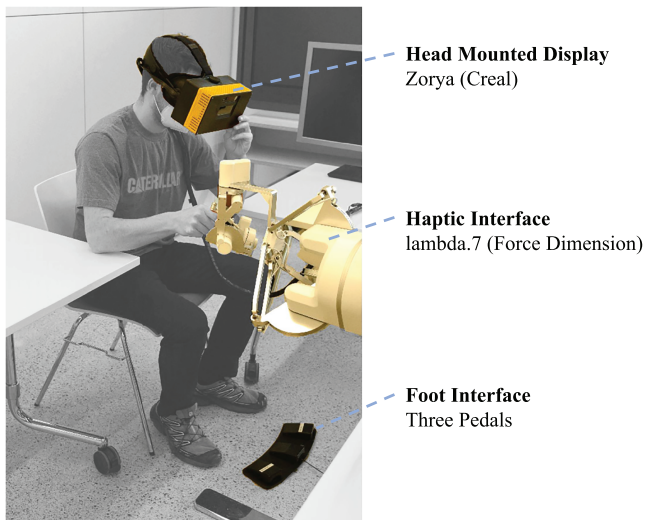Foot Interface
Three Pedals

FIGURE 10 Experimental setup

and neck and enabled the participant to keep the HMD in the right pose. This is important to ensure optimal visibility on the central light-field displays. An additional foam strip was attached to the HMD above the nose to prevent the display from fogging up. The left and right foot-pedals were used to start a trial or switch to a new configuration respectively.

2.4 | Experimental tasks

There were two types of experimental tasks: one 3D (three translations x,y,z) and four 1D (forward/backward translation z) interaction subtasks. The 1D tasks required depth *matching* in an abstract environment, while the 3D

tasks required *pick and place* procedures in a more realistic environment, thus allowing for a more gaming-like experience. The realistic environment in the *pick and place* task provided additional distance references, for example, through known objects surrounding the target position and shadows, and was suitable to make the participants familiar with the interaction principles and the HMD. In contrast, the 1D *matching* task emphasized the effects of light-field visualization by reducing the number of surrounding objects, shadows, and through using unknown objects with varying abstract shapes and simple textures.

In the *3D interaction task* (see Figure 11), a virtual apple which was controlled by the input device had to be picked up from a plate and moved to three target positions at different distances. Target positions were indicated by a transparent apple in the box. The perspective in VR was adjusted according to the tracked head movements of the subject. The input device could be moved freely in six degrees of freedom (DoF), but only translational movements were transferred to the VR while the orientation of the apples was set automatically to keep the task simple. There were three target positions in two different apple boxes. The first box was placed in the left half of the scene and is visualized in Figure 11. The second box was positioned in the right half and rotated 90° around the yaw-axis as well as tilted against roll-axis to the viewer.

The three target apples were located at the closest, central, and most distant area of the boxes each. Initially, subjects controlled a gray sphere in VR with the input device (compare Figure 11A). When reaching one of the apples on the plate (see Figure 11B), the respective apple replaced this sphere automatically overtaking the orientation of the respective target apple which appeared at the same time (see Figure 11B). The accuracy with which the pose of the target apples had to be matched (see Figure 11C,D) was set to 3 mm radial distance. After having matched the target position, the next trial was started.

The four *1D interaction tasks* are illustrated in Figures 12 and 13. Here, the VR perspective was locked, that is, the subjects' head motion was not tracked and transferred. This was done to exclude individual strategies of perspective switching to derive additional 3D information. The haptic interface could be moved freely, but only motions in z direction (forward/ backward motion in depth direction) were transferred to the VR. For ease of understanding, Figure 12 shows the different 1D interaction subtasks in a side view. Here, the objects shown as solid light lines are the planes controlled by the operator via the haptic interface. As mentioned before, only motions in z direction were possible. The reference or

FIGURE 11 The 3D interaction task: (A) initial position indicated by gray sphere, (B) picking of apple from plate, (C) approaching target position indicated by a transparent apple in the box, and (D) successfully matched target position
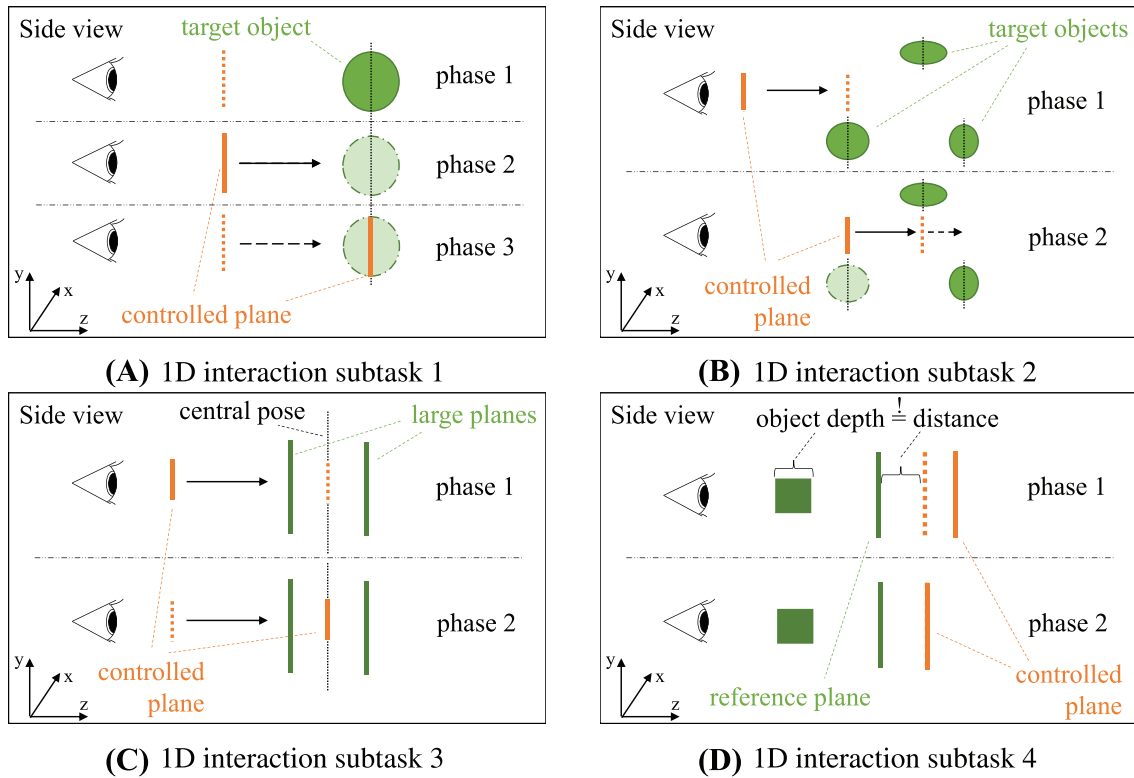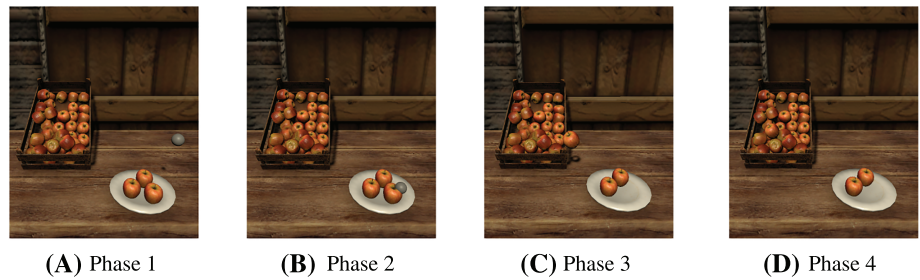
(A) Phase 1   **(B)** Phase 2   **(C)** Phase 3   **(D)** Phase 4

(A) 1D interaction subtask 1

(B) 1D interaction subtask 2

(C) 1D interaction subtask 3

(D) 1D interaction subtask 4

FIGURE 12 The 1D depth matching subtasks

target objects are shown in dark color. The dark dashed lines indicate the central z plane of the 3D objects (subtasks 1 and 2) or the central z plane between two large planes (subtask 3), respectively. The dotted light line marks the initial and the desired positions of the controlled object.

In subtask 1, shown in Figure 13A–C, the operator had to match the depth position of a small *controlled plane* (parallel to the x-y plane, shown on the left in Figure 12A, in light color) with the central depth plane of a stretched target object (visualized as the upper, dark sphere in Figure 12A). First, this target object appeared (phase 1) and disappeared (phase 2) again. Then, the light colored, *controlled plane* appeared at the initial depth plane (indicated by a dotted light colored line in Figure 12A). The participant had to remember the

position of the target object and moved the controlled plane in z direction in order to match the central depth plane of the reference sphere (phase 3).

Similarly, in subtask 2, the operator had to match the *controlled plane* with a set of five target objects as depicted in Figure 13D–F and simplified in Figure 13B. First, the central plane of the closest object had to be matched. Afterward, this object disappeared and the next closest object had to be matched. This routine had to be repeated until the most distant object had been processed.

In subtask 3 (compare Figure 12C), two large planes appeared at different distances in depth direction. These planes were positioned with a gap in between them in x direction (not visible in the side view, but from Figure 13G–I) such that the controlled plane could be

(A) Subtask 1 - phase 1   **(B)** Subtask 1 - phase 2   **(C)** Subtask 1 - phase 3

(D) Subtask 2 - phase 1   **(E)** Subtask 2 - phase 2a   **(F)** Subtask 2 - phase 2b

(G) Subtask 3 - phase 1a   **(H)** Subtask 3 - phase 1b   **(I)** Subtask 3 - phase 2

(J) Subtask 4 - phase 1a   **(K)** Subtask 4 - phase 1b   **(L)** Subtask 4 - phase 2   **(M)** Subtask 4 - phase 3
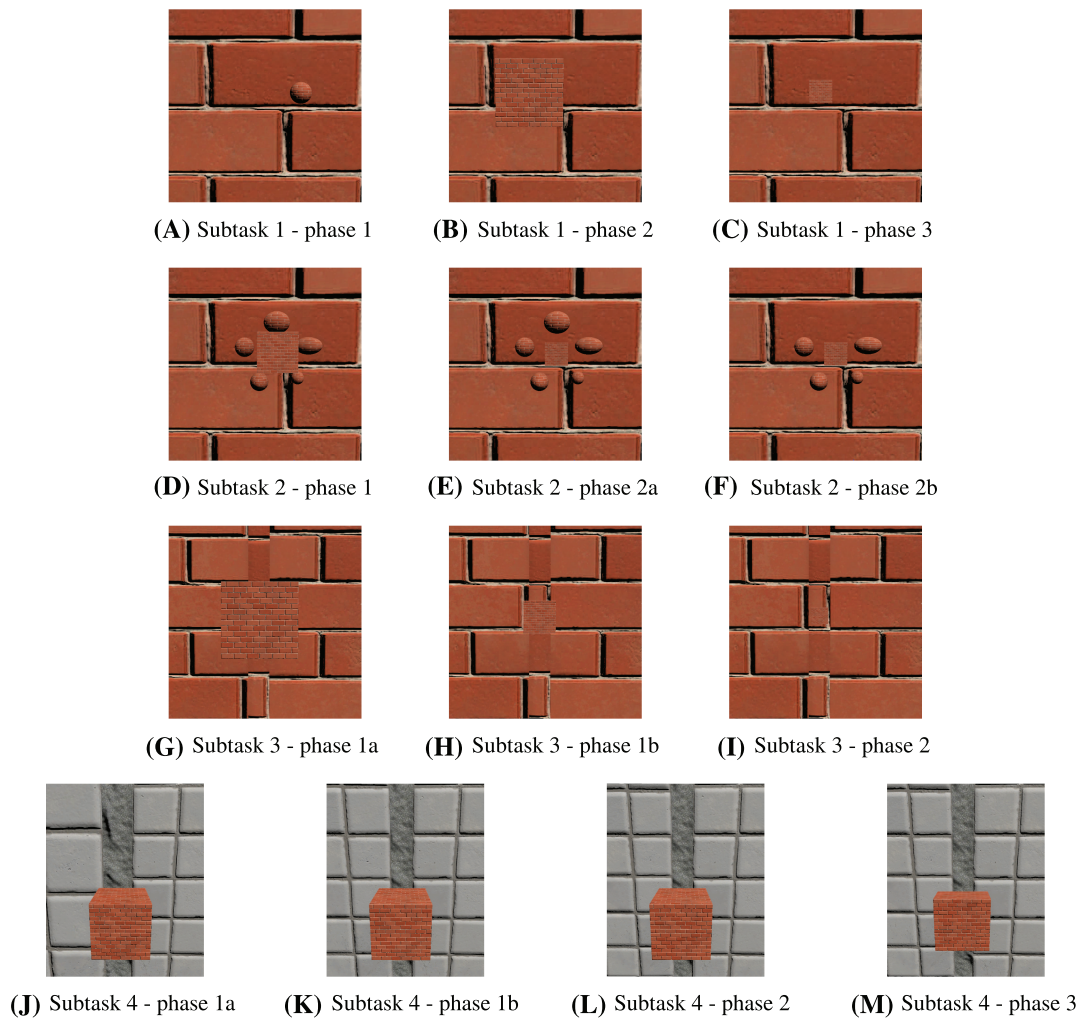
FIGURE 13   The 1D interaction tasks

moved between them without overlapping. The task was to position the *controlled plane* exactly in a central pose between the large planes in depth direction.

Figure 12D shows the 1D interaction subtask 4 with a cuboid and two large planes (again positioned left and right, as visible from Figure 13J–M). Instead of the small *controlled plane*, the test subject had to move the large left plane (presented in light color). This plane had to be positioned behind the right reference plane (presented in dark color) such that the z-distance of the planes was equal to the depth (width in z direction) of the object. The x and y position of the plane were fixed such that always both planes were visible and only the movement in z direction had to be controlled. This subtask 4 was repeated for five different object depths. With every new object, the position of the right plane was updated.

Four configurations (I to IV) were implemented for all 1D interaction subtasks, where configuration I and III as well as II and IV respectively were identical regarding object sizes, objects shapes and distances. While the

objects had unitary textures in configuration I (red brick wall) and II (white tiles), in configuration III and IV, the textures of the objects were mixed (red brick wall, white tiles and stony wall as presented in Figure 13M).

General design considerations

As analyzed in Banks et al.,[10] in conventional displays with fixed focal distance around 1.5 to 2 m, the VAC becomes very pronounced for vergence distances below 1 m. Thus, in this close range, the light-field displays are expected to bring the largest benefit. Therefore, this study investigates interaction tasks up to approximately 1 m corresponding to the arm reaching distances (with respect to the human eyes). Humans are able to perceive 3D from a vergence distance of approximately 0.2 m. Since for some visually impaired persons this distance can be higher, this study considered no object distances below 0.3 m. The tasks were classical *pick and place* or

matching tasks such that the target objects and the controlled planes were (initially) located in different depth layers. Therefore, the operator's eyes frequently had to switch between different depth layers. All tasks required a lot of refocusing and the 1D tasks required very high depth accuracy (1.5 mm).

2.5 | Experimental design

A within-subject design with *LIGHTFIELD* (*LF on* vs. *LF off*) as experimental factor was implemented, while the order of both conditions was counterbalanced across subjects. Moreover, in the 3D interaction task, each subject was assigned to one of three different target position orders (order 1: target 1, target 2, target 3; order 2: 2,3,1; order 3: 3,1,2). In the 1D interaction task, the configurations I to IV was systematically varied according to their texture such that one part of the test subjects had the order I, II, III, IV and the other III, IV, I, II. These configurations were systematically varied since they potentially interact with the light-field condition. Each of the experimental conditions started and ended with the 3D interaction task run, with the 1D interaction task block in between.

2.6 | Procedure

Table 1 shows the procedure of the study. The experimenter instructed the participants regarding study background, experimental conditions and tasks. Then, participants had to fill out an informed consent form and a demographic questionnaire. The following series of 3D interaction and 1D interaction tasks was performed twice in the light-field (*LF on*) and the non-light-field (*LF off*) condition (Table 1 (C) and (D)). After each trial, that is, six apples in the 3D interaction task and after each configuration (with four subtasks) in the 1D interaction task,

subjects orally answered the questions of an interim questionnaire, so they did not have to put off and readjust the HMD afterward.

After having finished an experimental condition, subjects put off the HMD and filled out a post-condition questionnaire. Before performing the 3D or 1D interaction task for the first time, subjects were familiarized with the tasks in a training trial. Before the first run of the 3D interaction task, in both conditions, the participant used the 3D environment to adjust the HMD correctly.

Instructions

The participants were told that the main performance criterion is the accuracy in the depth direction. It was recommended to use the arm rest for left and right arm and to hold the HMD with the left hand. When first attaching the HMD, the participants had to adjust the eye distance and to vary the vertical and horizontal HMD pose with respect to the head to find the optimal pose for the best visual performance of the HMD. This procedure was repeated when the HMD was reattached for the second block. Between all 1D interaction tasks, an object (Stanford Bunny) was displayed. The participants were requested to use this object to check the sharpness of visualization and thus if the HMD is still positioned correctly.

2.7 | Measures and statistical analysis

2.7.1 | Objective measures

The final position errors (distance of controlled object and target), motion path lengths (measured at input device) as well as completion times were recorded as objective measures.

TABLE 1 Experimental procedure

(A)	Instruction of participants
(B)	Informed consent and demographic questionnaire
(C)	1. 3D interaction task (run 1) 2. 1D interaction tasks (4 subtasks with 4 configurations) 3. 3D interaction task (run 2) 4. Post-condition questionnaire
(D)	1. 3D interaction task (run 1) 2. 1D interaction tasks (4 subtasks with 4 configurations) 3. 3D interaction task (run 2) 4. Post-condition questionnaire

2.7.2 | Subjective measures

In the interim questionnaire, participants subjectively rated the overall workload[35] (scale ranging from 1 = "very low" to 20 = "very high"), the degree to which they felt present in the VR (in %, 0% = "no experience of presence"; 100% = "like in the real world"), how confident they felt in performing the tasks (from 1 = "not at all" to 7 = "very confident"), the quality of depth perception (from 1 = "no depth perception at all" to 7 = "like in reality"), and finally, how well they could recognize objects (from 1 = "not at all" to 7 = "perfectly"), referring to the sharpness (exposure) of an object in front of the background or the objects in the close environment respectively as an indicator for the quality of the depth of field effect and the aid it provides to the user.

The post-condition questionnaire included the simulator sickness questionnaire (SSQ,[5] with nausea, disorientation and oculomotor symptom clusters) and the NASA Task Load Index (NASA-TLX,[36]) questionnaire.

2.7.3 | Statistical analysis

The objective measures of the 3D task were analyzed performing LIGHTFIELD * RUN * TARGET DISTANCE repeated measures ANOVAs (rmANOVA). For the measures in the 1D task LIGHTFIELD * CATEGORY * DISTANCE, rmANOVAs were conducted. A LIGHTFIELD * RUN rmANOVA was performed on the ratings of the interim questionnaire subsequent to the 3D tasks and accordingly, a LIGHTFIELD * CATEGORY rmANOVA for the 1D tasks. For all rmANOVAs, Greenhouse–Geisser corrections (GG.) were applied in case of non-sphericity. Paired t tests were conducted for comparing ratings in the *LF on* and *LF off* in the post-condition questionnaire.

3 | RESULTS

In the following analyzes, the objective data for the 1D interaction tasks from one subject (#17) had to be excluded, because the data of this participant was partially damaged due to a recording problem.

3.1 | Objective data

Table 2 summarizes the results in the performance data. The path length refers to the VR, that is, to the motion of the haptic interface scaled by 10.

TABLE 2 Results—Objective measures

	Light-Field OFF	Light-Field ON	Statistical Sign.
Completion Time [s]			
3D Interaction Task	5.93 (1.39)	5.93 (1.33)	ns.
1D Task—Subtask 1	4.86 (0.24)	4.88 (0.18)	ns.
1D Task—Subtask 2	2.84 (0.53)	2.62 (0.65)	$F(1,18) = 2.86; p = 0.05$ (1tt)
1D Task—Subtask 3	4.91 (0.17)	4.81 (0.26)	$F(1,18) = 3.86; p < 0.05$ (1tt)
1D Task—Subtask 4	4.94 (0.13)	4.91 (0.17)	ns.
Path Length [m]			
3D Interaction Task	1.56 (0.26)	1.59 (0.31)	ns.
1D Task—Subtask 1	0.55 (0.24)	0.56 (0.28)	ns.
1D Task—Subtask 2	0.20 (0.09)	0.19 (0.10)	ns.
1D Task—Subtask 3	0.76 (0.15)	0.72 (0.11)	$F(1,18) = 3.82; p < 0.05$ (1tt)
1D Task—Subtask 4	0.34 (0.11)	0.33 (0.12)	ns.
Final Position Error [m]			
3D Interaction Task	0.019 (0.001)	0.020 (0.001)	ns.
1D Task—Subtask 1	0.151 (0.107)	0.143 (0.161)	ns.
1D Task—Subtask 2	0.013 (0.015)	0.015 (0.034)	ns.
1D Task—Subtask 3	0.080 (0.078)	0.073 (0.115)	ns.
1D Task—Subtask 4	0.092 (0.062)	0.085 (0.056)	ns.

3.1.1 | The 3D interaction task

Analyzing the performance metrics (completion time, path length, and final position error) revealed no significant main effects of LIGHTFIELD. Yet, significant main effects of RUN were found for completion times [$F(1,18) = 10.81$; $p < 0.01$] and path lengths [$F(1,18) = 13.14$; $p < 0.01$], indicating simple learning effects, that is, subjects performed better in the second run. Not surprisingly, for both variables, TARGET DISTANCE reached significance [Completion time: $F(1,18) = 5.60$; $p < 0.001$; Path length: $F(2.97,53.44) = 7.17$, GG.; $p < 0.001$]. Although, a significant LIGHTFIELD*TARGET DISTANCE interaction effect [$F(5,90) = 2.39$; $p < 0.05$] was evident for completion times, no significant effect of LIGHTFIELD was found for any distance.

3.1.2 | The 1D interaction task

RmANOVA on completion times across all subtasks showed that subjects completed the 1D interaction task faster when light-field was activated [$F(1,18)=3.6$; $p < 0.05$, one-tailed testing]. No such overall effect was found for path lengths and final position error. Analyzing the subtasks separately, revealed that the positive effect of LIGHTFIELD on completion time was most evident in subtasks 2 [$F(1,18)=2.9$; $p = 0.05$, one-tailed testing] and 3 [$F(1,18)=3.9$; $p < 0.05$, one-tailed testing]. In subtask 2, a significant LIGHTFIELD * CATEGORY interaction effect [$F(3,54)=3.3$; $p < 0.05$] furthermore indicated the

strongest effects of LIGHTFIELD for category II, that is, when the set of spheres was more distant (0.64–0.75 m) and had the same texture. Here, subjects performed significantly worse in the *LF off* condition when compared to the other categories (all ps < 0.05). Also, for subtask 3 a significant LIGHTFIELD * CATEGORY interaction effect [$F(3,54)=3.0$; $p < 0.05$] emerged, showing that the positive effect of LIGHTFIELD was only significant in category I ($p < 0.05$), that is, when target planes were more distant (1.15 m) and had the same texture.

Additionally, a significant main effect of LIGHTFIELD on path lenghts was found in subtask 3 [$F(1,18)=3.8$; $p < 0.05$, one-tailed testing], that is, path lengths were shorter in the *LF on* compared to the *LF off* condition.

3.2 | Subjective data

3.2.1 | Interim questionnaire

Table 3 summarizes the results of the interim questionnaire for both task types and experimental conditions.

The 3D interaction task

RmANOVA was performed on the different items of the questionnaire after the experimental runs. Only a marginally significant trend for LIGHTFIELD was found for "depth perception" [$F(1,19)=2.9$; $p = 0.05$, one-tailed

TABLE 3 Results—Interim questionnaire

	Light-Field OFF	Light-Field ON	Statistical Sign.
Overall Workload			
3D Interaction Task	3.78 (2.61)	4.48 (3.20)	ns.
1D Interaction Task	9.36 (3.91)	9.96 (4.33)	ns.
Sense of Presence			
3D Interaction Task	76.43 (15.13)	77.50 (15.54)	ns.
1D Interaction Task	64.40 (16.04)	65.01 (17.52)	ns.
Confidence			
3D Interaction Task	6.70 (0.47)	6.58 (0.49)	ns.
1D Interaction Task	4.60 (1.01)	4.96 (0.82)	F(1,19) = 4.32; $p < 0.05$ (1tt)
Depth Perception			
3D Interaction Task	6.15 (0.61)	6.38 (0.63)	F(1,19) = 2.85; $p = 0.05$ (1tt)
1D Interaction Task	4.88 (0.86)	5.26 (1.08)	F(1,19) = 5.89; $p < 0.05$ (2tt)
Object Detection			
3D Interaction Task	6.15 (0.54)	6.20 (0.77)	ns.
1D Interaction Task	5.11 (1.21)	5.68 (1.09)	F(1,19) = 6.83; $p < 0.05$ (2tt)

testing], that is, depth perception was rated as being better with light-field. A significant RUN main effect [$F(1,19)=7.0$; $p<.05$, two-tailed testing] revealed better ratings in the second compared to the first run.

The 1D interaction task

rmANOVA yielded several significant main effects of LIGHTFIELD. Subjects rated their 'confidence' being higher in *LF on* compared to *LF off* [$F(1,19)=4.3$; $p < 0.05$, one-tailed testing]. A significant CATEGORY main effect for this item [$F(3,57) = 7.2$; $p < 0.01$, two-tailed testing], showed that ratings were generally higher for categories III and IV compared to I and II, that is, subjects were more confident when textures of controlled and target objects differed. Moreover, "depth perception" and "object detection" were rated better when light-field was activated [$F(1,19) = 5.9$; $p < 0.05$ and $F(1,19) = 6.8$;

both $ps < 0.05$; both two-tailed testing]. For both items, the same CATEGORY main effect occurred as for "confidence" [$F(3,57) = 5.9$ or 4.1, respectively; both $ps < 0.05$, two-tailed testing].

3.2.2 | Post-condition questionnaire

In a first step, the sum scores of both questionnaires were calculated and paired t-tests were performed on these scores. While no significant effect was found for the SSQ total score, the NASA-TLX sum score tended to be lower in the *LF on* condition ($t(19)=1.4$; $p < 0.10$, one-tailed testing). Comparing the individual items of the SSQ revealed marginally significant effects for "eye strain" ($t(19) = 1.5$; $p < 0.10$, one-tailed testing) and "fullness of head" ($t(19) = 1.7$; $p = 0.05$, one-tailed testing) and a significant effect of "difficulty concentrating" ($t(19) = 1.8$; $p < 0.05$, one-tailed testing). A significant

TABLE 4 Results—Post-condition questionnaire (simulator sickness cluster: n = nausea, o = oculomotor, d = disorientation)

	Light-Field OFF	Light-Field ON	Statistical Sign.	Cluster
Simulator Sickness				
(SSQ Total Score)	258.0 (191.6)	226.2 (211.4)	ns.	
General discomfort	1.00 (0.80)	1.05 (0.76)	ns.	n, o
Fatigue	0.50 (0.69)	0.65 (0.75)	ns.	o
Headache	0.05 (0.22)	0.10 (0.31)	ns.	o
Eye strain	0.55 (0.69)	0.35 (0.49)	$t(19) = 1.45$; $p < 0.10$ (1tt)	o
Difficulty Focussing	0.40 (0.60)	0.40 (0.82)	ns.	o,d
Salivation	0.05 (0.22)	0.10 (0.31)	ns.	n
Sweating	0.20 (0.52)	0.15 (0.37)	ns.	n
Nausea	0.15 (0.37)	0.10 (0.31)	ns.	n,d
Difficulty Concentrating	0.30 (0.47)	0.15 (0.37)	$t(19) = 1.83$; $p < 0.05$ (1tt)	n,o
Fullness of head	0.75 (0.85)	0.55 (0.89)	$t(19) = 1.71$; $p = 0.05$ (1tt)	d
Blurred vision	0.30 (0.57)	0.20 (0.52)	ns.	o,d
Dizziness (eyes open)	0.15 (0.37)	0.15 (0.37)	ns.	d
Dizziness (eyes closed)	0.25 (0.44)	0.20 (0.52)	ns.	d
Vertigo	0.00 (0.0)	0.00 (0.0)	ns.	d
Stomach awareness	0.15 (0.49)	0.10 (0.31)	ns.	n
Burping	0.00 (0.0)	0.05 (0.22)	ns.	n
Workload				
(Raw TLX Sum Score)	45.65 (16.37)	42.60 (18.68)	$t(19) = 1.35$; $p < 0.10$ (1tt)	
Mental Demand	8.45 (3.99)	7.95 (4.63)	ns.	
Physical Demand	6.20 (4.23)	6.30 (4.59)	ns.	
Temporal Demand	8.55 (3.53)	7.75 (3.71)	ns.	
Performance	8.15 (3.30)	7.35 (3.62)	$t(19) = 1.85$; $p < 0.05$ (1tt)	
Effort	9.35 (4.39)	8.70 (4.68)	ns.	
Frustration	4.95 (3.32)	4.55 (3.73)	ns.	

effect for the NASA-TLX item "performance" (M_{LF}=7.35; M_{noLF} = 8.15; 1 = "perfect"; 20 = "failure"; $t(19)$ = 1.8; $p < 0.05$, one-tailed testing) also indicated that participants rated their performance to be better when light-field was active. The results of the post-condition questionnaire are summarized in Table 4.

4 | DISCUSSION

In the following, the results are first discussed with respect to the hypotheses. Then, general interpretations are noted.

4.1 | H1

The SSQ comprises the three different symptom clusters oculomotor, disorientation and nausea (o, d, n in Table 4[5]). The nausea cluster describes somatic effects like stomach awareness. Here, we found no significant effect of light-field besides the item "difficulty concentrating" which is also part of the oculomotor cluster. The oculomotor cluster describes the effort of the eye system and the disorientation cluster relates to the central nervous processing of visual information including the vestibular system. "Eye strain" and "difficulty concentrating" are items of the oculomotor cluster and thus affected for example by the additional muscular effort of the eye when no light-field information is provided by the display. The factor "fullness of head" (describing a pressure from the inside) belongs to the disorientation cluster. The results indicate an increased effort of the central nervous system in case of *LF off* since the brain has to process inconsistent visual information (VAC). Thus, H1 can be considered as confirmed for distances below 1.15 m which is the range relevant for interaction tasks.

4.2 | H2

The significant improvement of "depth perception" and "object detection" (1D interaction task) confirms that the participants were able to perceive the light-field effects, although only two test subjects were aware that in the two experimental sections light-field was activated or deactivated, respectively. The significant improvement of confidence, ease of concentration as well as the lower NASA-TLX sum score indicate that light-field displays improves the user experience and reduce workload, thus confirming H2. This might result from the prevention of the VAC as well as from the improved, more natural depth presentation with depth of field effects.

The positive effects of *LF on* on the visual effort becomes visible from the subjective questionnaire data. It can be assumed that the visual effort affects the performance of the user mainly after prolonged use of HMDs. The found significant results in the performance data of this study can be assumed to originate from the improved depth perception through light-field resulting especially from the effects of depth sharpness. Thus, the objective results are linked rather to H2 than H1. In subtasks 1 and 4 of the 1D-task, no significant effects were found, probably, since both tasks required memorizing the position or depth of the reference object which might be more challenging than depth estimation itself. Still, the reduced completion time in subtask 2 strengthens the results regarding subjective rating of improved depth estimation which might have facilitated faster matching of the target position. Subtask 3 had a similar design as subtask 4, but the reference objects were located close to the controlled object, and thus, memory effects do not play a role. The reduced completion time and path length indicate that the user perceived the target position more easily such that less or only lower overshoots over the target position occurred. With *LF off*, in a completely sharp view, the users seemingly needed more back and forth motions to estimate the depth through stereoscopy. As expected, the positive effects were overall more evident for configurations with unitary textures. Positive effects of light-field were found in the performance data mainly for more distant objects. This might be due to the fact that stereoscopy eases depth estimation especially for closer objects.

Overall, the results show that the light-field HMD brings clear benefits to the user in terms of physical (reduced eye strain, fullness of head) and mental use comfort (improved concentration, estimated performance, and reduced times). The visual effort is reduced through the avoidance of the VAC, while the depth perception is in addition improved through the depth of field which provides a more natural distinction between in-focus foreground objects and a distant blurry background. With these benefits, the light-field technology promises to increase the use comfort and consequently the interaction performance in VR training as well as in telerobotic systems reaching from industrial to surgical setups. In the same way, AR applications for navigation or assembly and maintenance instructions or the Metaverse can be enhanced through light-field displays.

4.3 | Limitations

Overall, the study lasted in average 1.5 h for the participants. During the first block with training phases, the HMD was worn for 30 to 40 min and for another 20–25

min during the second block. Although the 1D interaction task was designed to be highly demanding and stressful for the eyes, the participants reported comparably very low scores in the simulator sickness questionnaire. This indicates that the problem of the VAC might become stronger after longer use of non-light-field displays or are already reduced through the high resolution of the light-field display.

A major challenge of the study was the verification that the user wears the display correctly such that the optimal perception is achieved. Still, the results on depth perception and object recognition hint that the display was worn correctly. This can be further improved in the future through the integration of eye tracking technologies.

Despite the adjustment possibilities of the HMD, it was beneficial to additionally hold it with one hand, due to its weight (caused by its prototyped status).

Due to the complexity of the study and the hardware setup, two persons had to be inside the same room. Thus, according to the DLR Covid-19 rules at the time of the experiment, the principal investigator and the participant had to wear FFP2 masks. This clearly affected the overall usability of the HMD since additional cushions had to be added which additionally aggravated breathing.

5 | CONCLUSION

This work presented the first statistical analysis of benefits of light-field head mounted displays on visual effort and the first study involving kinesthetic interaction tasks. The study was performed in a well determined environment under consistent conditions. Since the same hardware was applied for the non-light-field and light-field conditions, a potential bias of the user related to novel hardware could be prevented.

The results show that the light-field HMD brings clear benefits to the user in terms of physical use comfort (reduced eye strain and fullness of head) and mental use comfort (improved concentration, estimated performance, and reduced times). Firstly the light-field technology reduces the visual effort through solving the VAC, secondly the depth perception is improved through the depth of field which provides a more natural distinction between in-focus foreground objects and a distant blurry background.

In future work, the light-field HMDs will also be integrated into a robotic tele-operation setup to evaluate the technology with real-time non-virtual content.

ORCID
Michael Panzirsch ![ORCID icon] https://orcid.org/0000-0002-0647-7147

REFERENCES
1. Da Vinci L, Amoretti C, et al. Trattato Della Pittura di Lionardo da Vinci, vol. 201: Soc. Tipogr. de Classici Italiani; 1804.
2. Wheatstone C. Xviii. Contributions to the physiology of vision. Part the first. on some remarkable, and hitherto unobserved, phenomena of binocular vision. Philos Trans R Soc Lond. 1838; 128:371–394.
3. Julesz B. Binocular depth perception of computer-generated patterns. Bell Syst Techn J. 1960;39(5):1125–1162.
4. Caserman P, Garcia-Agundez A, Zerban AG, Göbel S. Cybersickness in current-generation virtual reality head-mounted displays: systematic review and outlook. Virtual Reality. 2021;25:1153–1170.
5. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int J Aviat Psychol. 1993;3(3):203–220.
6. Bando T, Iijima A, Yano S. Visual fatigue caused by stereoscopic images and the search for the requirement to prevent them: a review. Displays. 2012;33(2):76–83.
7. Hoffman DM, Girshick AR, Akeley K, Banks MS. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. J Vis. 2008;8(3):33–33.
8. Lambooij M, Fortuin M, Heynderickx I, IJsselsteijn W. Visual discomfort and visual fatigue of stereoscopic displays: a review. J Imaging Science Technol. 2009;53(3):30201–1.
9. Banks MS. I3. 1: invited paper: the importance of focus cues in stereo 3d displays. Sid symposium digest of technical papers: Wiley Online Library, 2015. p. 282–285.
10. Banks MS, Kim J, Shibata T. insight into vergence/accommodation mismatch. head-and Helmet-Mounted Displays XVIII: design and applications. Int Soc Opt Photon; 2013;8735:873509.
11. Banks MS, Sprague WW, Schmoll J, Parnell JA, Love GD. Why do animal eyes have pupils of different shapes? Sci Adv. 2015; 1(7):e1500391.
12. Kramida G. Resolving the vergence-accommodation conflict in head-mounted displays. IEEE Trans Vis Comput Graph. 2015; 22(7):1912–1931.
13. Yamaguchi M. Light-field and holographic three-dimensional displays. JOSA A. 2016;33(12):2348–2364.
14. Zhan T, Xiong J, Zou J, Wu S-T. Multifocal displays: review and prospect. PhotoniX. 2020;1:1–31.
15. Park J-H, Kim S-B. Optical see-through holographic near-eye-display with eyebox steering and depth of field control. Opt Express. 2018;26(21):27,076–27,088.
16. Tan G, Zhan T, Lee Y-H, Xiong J, Wu S-T. Polarization-multiplexed multiplane display. Opt letters. 2018;43(22):5651–5654.
17. Zhan T, Zou J, Lu M, Chen E, Wu S-T. Wavelength-multiplexed multi-focal-plane seethrough near-eye displays. Opt Express. 2019;27(20):27507–27513.
18. Konrad R, Padmanaban N, Molner K, Cooper EA, Wetzstein G. Accommodation-invariant computational near-eye displays. ACM Trans Graph (TOG). 2017;36(4):1–12.

19. Rathinavel K, Wang H, Blate A, Fuchs H. An extended depth-at-field volumetric near-eye augmented reality display. IEEE Trans Vis Comput Graph. 2018;24(11):2857–2866.

20. Dunn D, Tippets C, Torell K, Kellnhofer P, Akşit K, Didyk P, et al. Wide field of view varifocal near-eye display using see-through deformable membrane mirrors. IEEE Trans Vis Comput Graph. 2017;23(4):1322–1331.

21. Lai Y-M, Hsu C-H. Refocusing supports of panorama light-field images in head-mounted virtual reality. *Proceedings of the 3rd international workshop on multimedia alternate realities*; 2018. p. 15–20.

22. Raimbaud P, Lopez MSA, Figueroa P, Hernandez JT. Influence of depth cues on eye tracking depth measurement in augmented reality using the magicleap device. *2020 IEEE conference on virtual reality and 3d user interfaces abstracts and workshops (VRW)*, IEEE; 2020. p. 210–214.

23. Lin G, Panigrahi T, Womack J, Ponda DJ, Kotipalli P, Starner T. Comparing order picking guidance with Microsoft hololens, magic leap, google glass xe and paper. *Proceedings of the 22nd international workshop on mobile computing systems and applications*; 2021. p. 133–139.

24. Xu C. Nreal: ready-to-wear mixed reality glasses. Spie avr21 Industry Talks II, 2021. p. 1176409.

25. Waveguide optics. https://www.vuzix.com/technology/optics-waveguides-display-engines; 2021.

26. Dewald DS, Evans AT, Welch N, Gross A, Hill G. 8-1: invited paper: the avegant glyph: optical design considerations and approach to near-eye display. Sid Symposium Digest of Technical Papers: Wiley Online Library, 2016. p. 69–71.

27. Osmanis I. Lightspace technologies: eye accommodation realized by multifocal projection. SPIE AVR21 Industry Talks II, 2021. p. 1176413.

28. Milne DF. Vividq: indistinguishable from reality: overcoming key challenges in highly realistic holographic display for ar and beyond. Spie avr21 industry talks II: International Society for Optics and Photonics, 2021. p. 1176405.

29. Real time computational holography. https://www.cyvision.com/technology; 2021.

30. Leister N, Schwerdtner A, Fütterer G, Buschbeck S, Olaya J-C, Flon S. Full-color interactive holographic projection system for large 3d scene reconstruction. Emerging Liquid Crystal Technologies III: International Society for Optics and Photonics, 2008. p. 69110V.

31. Sluka T. Near-eye sequential light-field projector with correct monocular depth cues. *European Patent EP3542206B1*; 2017.

32. Battisti F, Carli M, Le Callet P. A study on the impact of visualization techniques on light field perception. *2018 26th European Signal Processing Conference (EUSIPCO)*, IEEE; 2018. p. 2155–2159.

33. Gotsch D, Zhang X, Carrascal JP, Vertegaal R. Holoflex: a flexible light-field smartphone with a microlens array and a p-oled touchscreen. *Proceedings of the 29th annual symposium on user interface software and technology*; 2016. p. 69–79.

34. Shidanshidi H, Safaei F, Li W. Estimation of signal distortion using effective sampling density for light field-based free viewpoint video. IEEE Trans Multimed. 2015;17(10):1677–1693.

35. Vidulich MA, Tsang PS. Absolute magnitude estimation and relative judgement approaches to subjective workload assessment. Proceedings of the Human Factors Society Annual Meeting. Los Angeles, CA: SAGE Publications Sage CA, 1987. p. 1057–1061.

36. Hart SG, Staveland LE. Development of nasa-tlx (task load index): results of empirical and theoretical research. Advances in Psychology: Elsevier, 1988. p. 139–183.

AUTHOR BIOGRAPHIES

Michael Panzirsch received his diploma in mechanical engineering from the Technische Universität München in 2010. Since then he is with the Department for Analysis and Control of Advanced Robotic Systems of the German Aerospace Center (DLR) in Oberpfaffenhofen as a research associate. He finished his PhD thesis on passivity-based multilateral control for delayed teleoperation at the Polytechnical University of Madrid (UPM) in October 2018. His main areas of research interests are teleoperation of stationary and mobile robots, haptics and healthcare robotics. Currently, he is working on topics as Haptic Augmentation, Model-Augmented Teleoperation and Shared Control.

Bernhard Weber received his Dipl.-Psych. and PhD degree at the University of Würzburg, Germany, in 2004 and 2008, respectively. From 2008–2010 he was with the German Aerospace Center (DLR), Institute of Flight Guidance, Brunswick, Germany, and since 2010 at the DLR Institute of Robotics and Mechatronics, Wessling, Germany, as a human factors expert. His main research interests are human factors in telerobotic systems, evaluation of haptic interaction technology, and sensorimotor performance under conditions of microgravity.

Nicolai Bechtel received his Master of Science in Computational Engineering from the University of Applied Sciences Munich in 2018. Since then, he has been conducting research in the field of haptics and virtual reality as a research assistant at the Center for Robotics and Mechatronics of the German Aerospace Center (DLR) in Oberpfaffenhofen. His research focuses on haptics, multi-dynamic simulations, and development of virtual reality environments. He is currently working on topics such as Multi-Contact Haptics and Model-Based Teleoperation.

Nicole Grabner finished her Bachelors in Robotics and Mechatronics at the FH Technikum Wien in 2019. She is currently enrolled in the Master-Program Robotics, Cognition and Intelligence at the Technische Universität München (TUM), while working at the German Aerospace Center (DLR) in Oberpfaffenhofen as a working student. There, she has been mainly working on VR-Visualization Systems, Eye Tracking and Communication Protocols with haptic feedback devices.

Martin Lingenauber is researcher and project leader in the field of robot vision at the DLR Institute of Robotics and Mechatronics. In his research, he combines his space engineering background with his interest in computer vision and robotics. Therefore, his work is mostly related to space robotic topics, mostly with a focus on new sensor concepts, such as light field cameras a.k.a. plenoptic cameras, and depth perception. Besides his research, he's also interested in software engineering topics for research and he likes creating or enhancing experimental setups.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.