Supplementary Material: Reproducing Reality with a High-Dynamic-Range Multi-Focal Stereo Display

This is a supplementary document for the paper "Reproducing Reality with a High-Dynamic-Range Multi-Focal Stereo Display" (DOI: 10.1145/3478513.3480513). Here we present (i) a detailed description of each component of our display setup with CAD drawings and (ii) an analysis of individual performance for our visual Turing test.

1 Additional details of the HDR-MF-S display apparatus

This section describes the details of the display construction, including CAD drawings. Note that the drawings are given as accurate as possible. However, during the actual construction of the setup, there were many custom parts that we redesigned and implemented differently. Therefore, some of the components in the CAD drawings might be different than the photographs of the final setup. Fig. 1 shows the entire display system from different perspectives, i.e., top-front view, top-back view, top-front view with open covers, top-back view with open covers. The setup comprises four HDR display units (Fig. 3), a combiner optics (Fig. 4,) and a camera gantry (Fig. 6) as shown in Fig. 1. The system components are described in detail in the following subsections.

1.1 The HDR displays

The details of the HDR display units are discussed in Sec. 3.2 of the main paper. Here, we give the details of the construction in terms of the dimensions and build in the CAD drawings. Fig. 2 shows two pairs of HDR displays (bottom and top). Each pair is a stereoscopic display which creates one focal depth. The bottom and top stereoscopic display pairs are placed at variable distances from the eye, controlled by sliding each unit on its aluminium profile. The distance from the eye is measured by considering the light path that is reflected by the beam-splitter and the mirror. In the top view of the setup in Fig. 2, one may notice that the display units are identical but not symmetrical. Therefore, the two units on the right are shifted away from the observer so that he projected images for the left and right eye are aligned. A close-up CAD drawing of a single HDR unit is shown in Fig. 3.

1.2 The combiner optics

The setup is built to allow the participant to observe the real scene and the virtual images from the same view port via combiner optics as shown in Fig. 4. The front focal plane images are generated by the bottom stereoscopic HDR pair and the images pass through (50R/50T, weidner-glas.de) beam-splitters and reflected by the (70R/30T, Edmund Optics, 64-409) beam-splitters. The rear focal plane images are generated by the top stereoscopic HDR pair and the images first reflected by the first-surface reflection mirrors (weidner-glas.de), the (50R/50T) beam-splitters, then reflected by the (70R/30T) beam-splitters. The beam-splitters are mounted on a custom designed and 3D printer mounts. The mounts for the (70R/30T) beam-splitters are fit into the symmetrical optical mirror mounts (Thorlabs KM200S and KM200SL) as shown in Fig. 5.

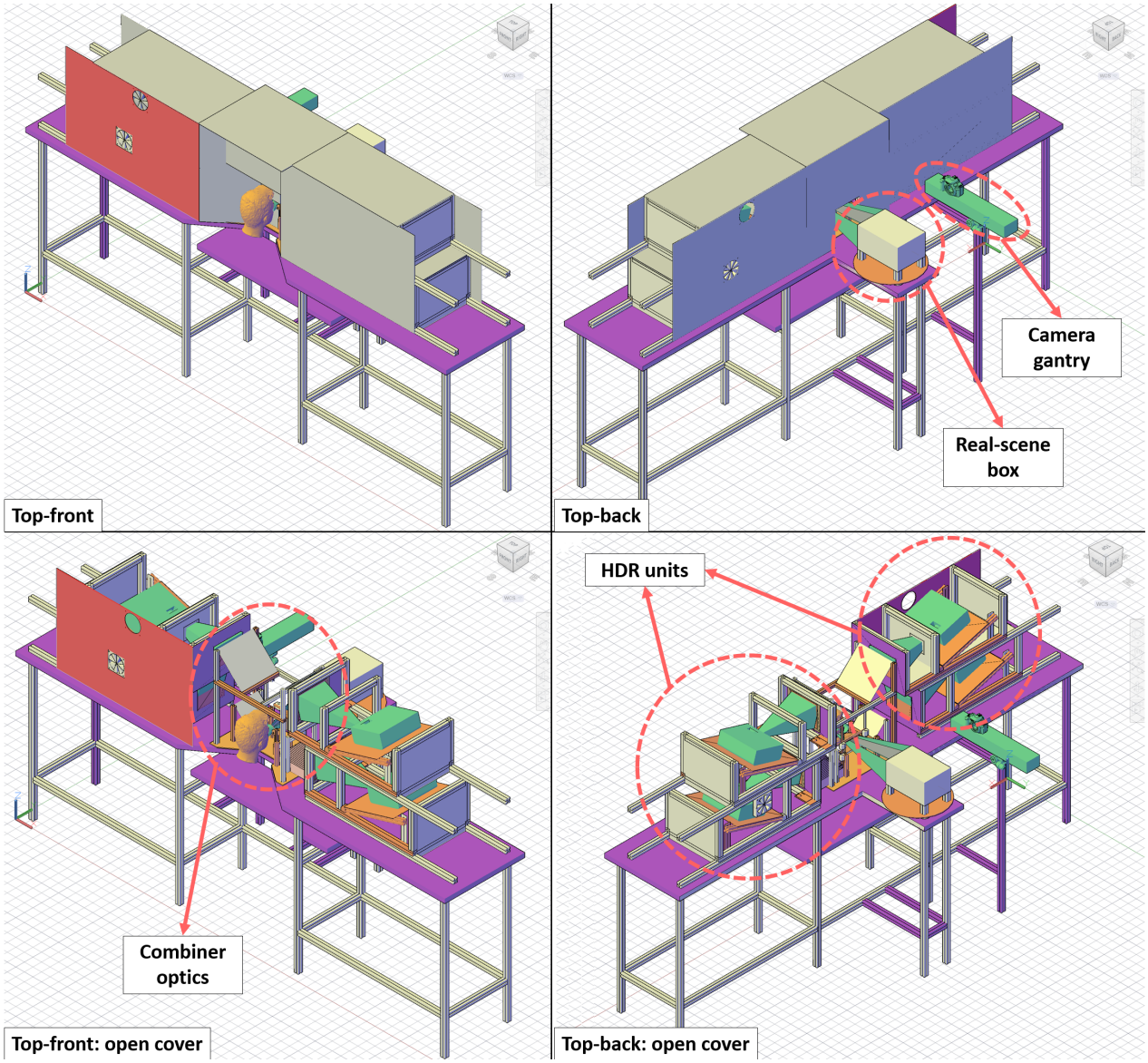

Figure 1: The CAD drawings of the entire system.

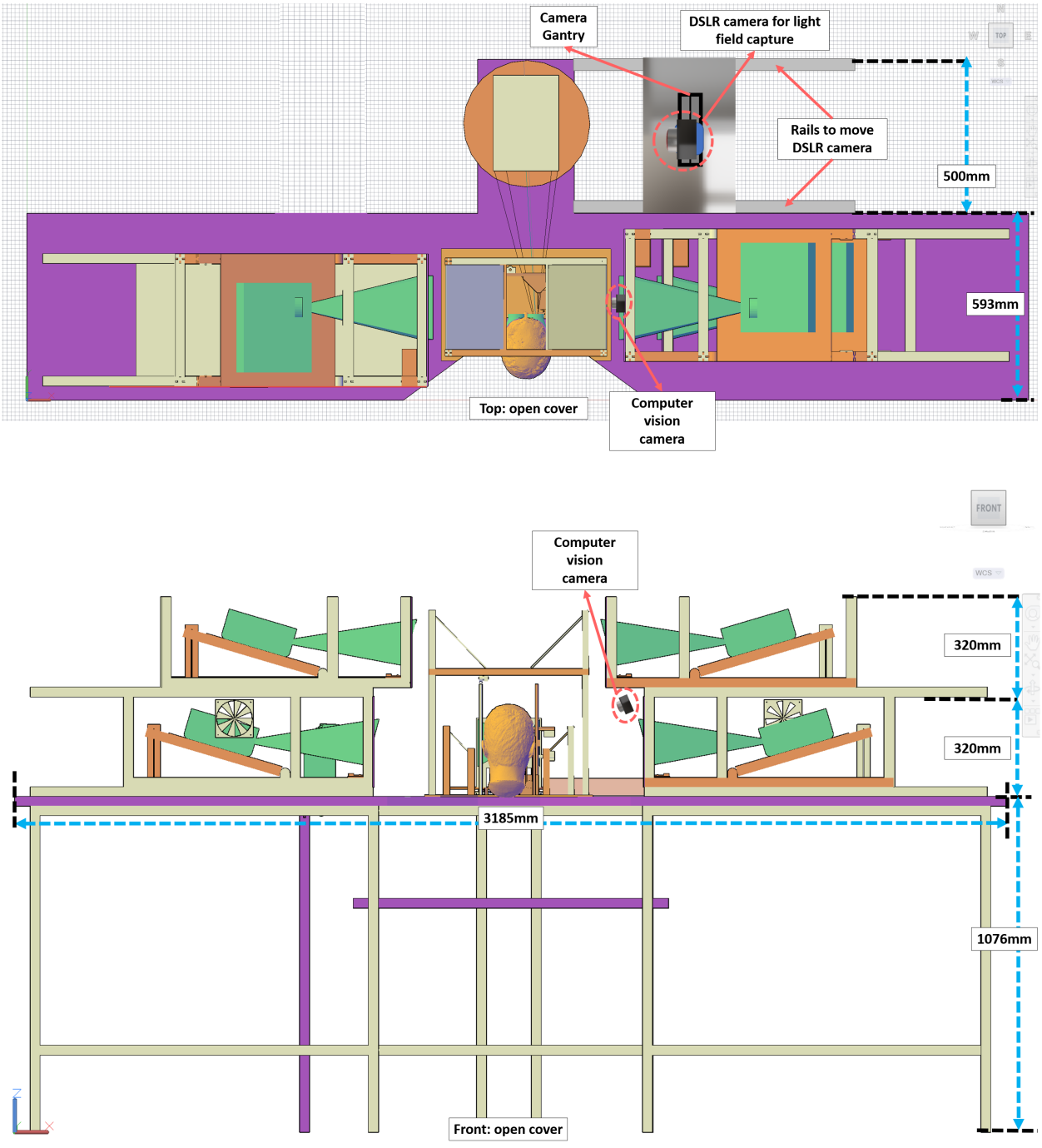

Figure 2: The CAD drawing of the HDR pairs from the top and front views with covers open. (Note that we have removed the camera gantry shown in Fig. 1 in this figure. Instead, we drew a top view of the DSLR camera, camera gantry, and the platform that we can move towards and away from the real-scene box. This reflects the actual setup.)

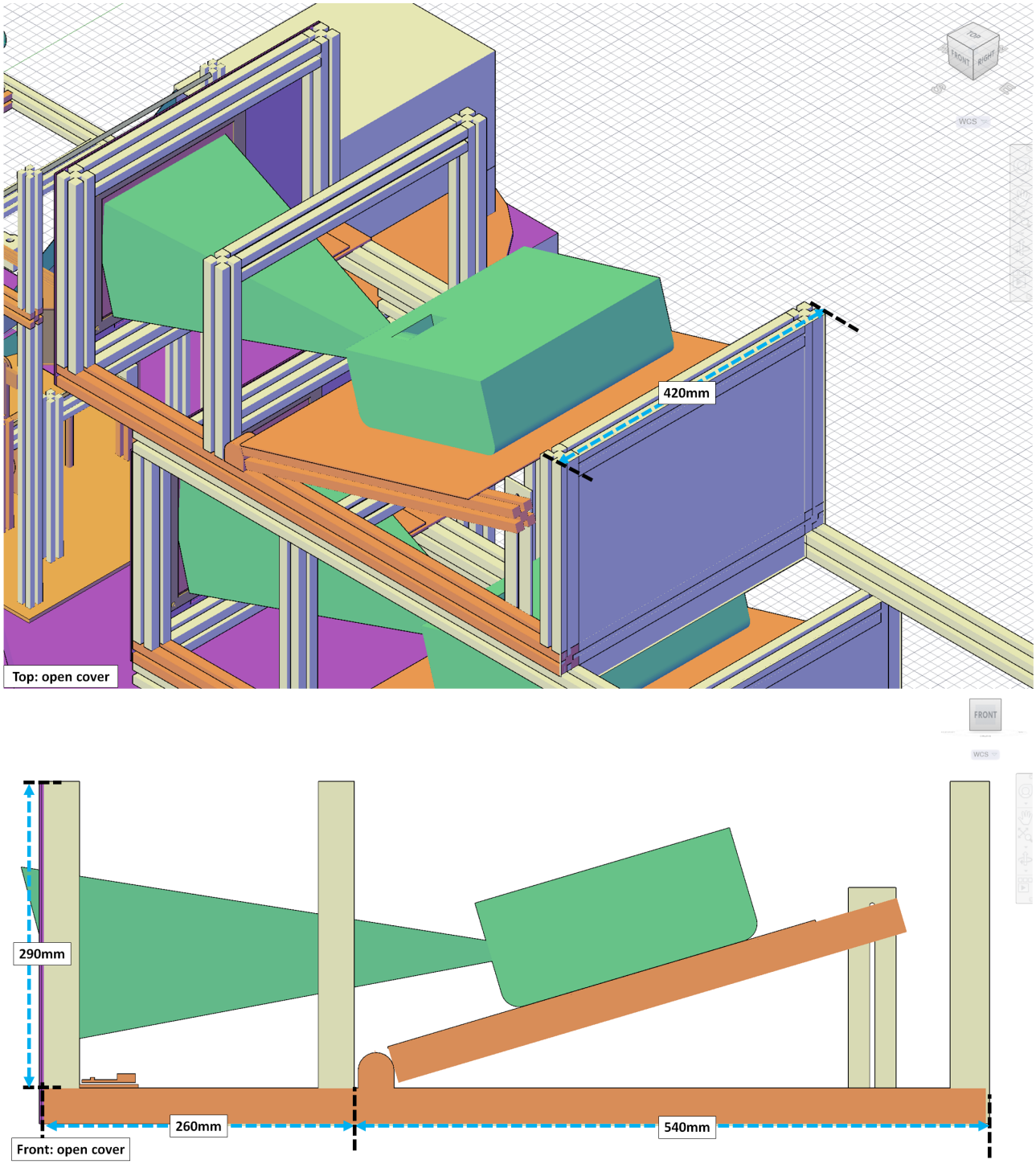

Figure 3: The CAD drawings of a single HDR unit from top-front view and front view with open covers. The green pyramids are modelling the light beams emanating from the projectors.

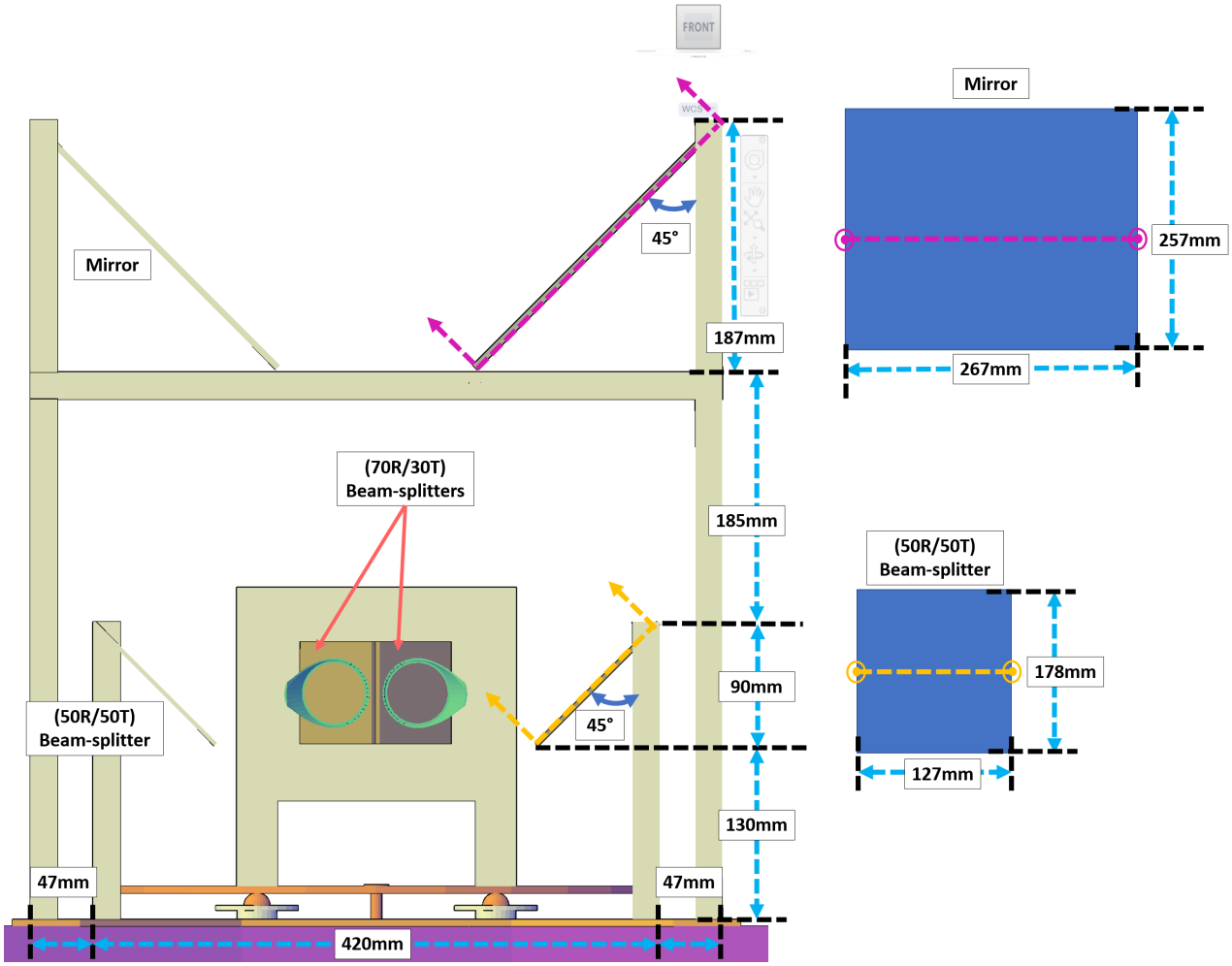

Figure 4: The CAD drawing of the view port and the combiner optics.

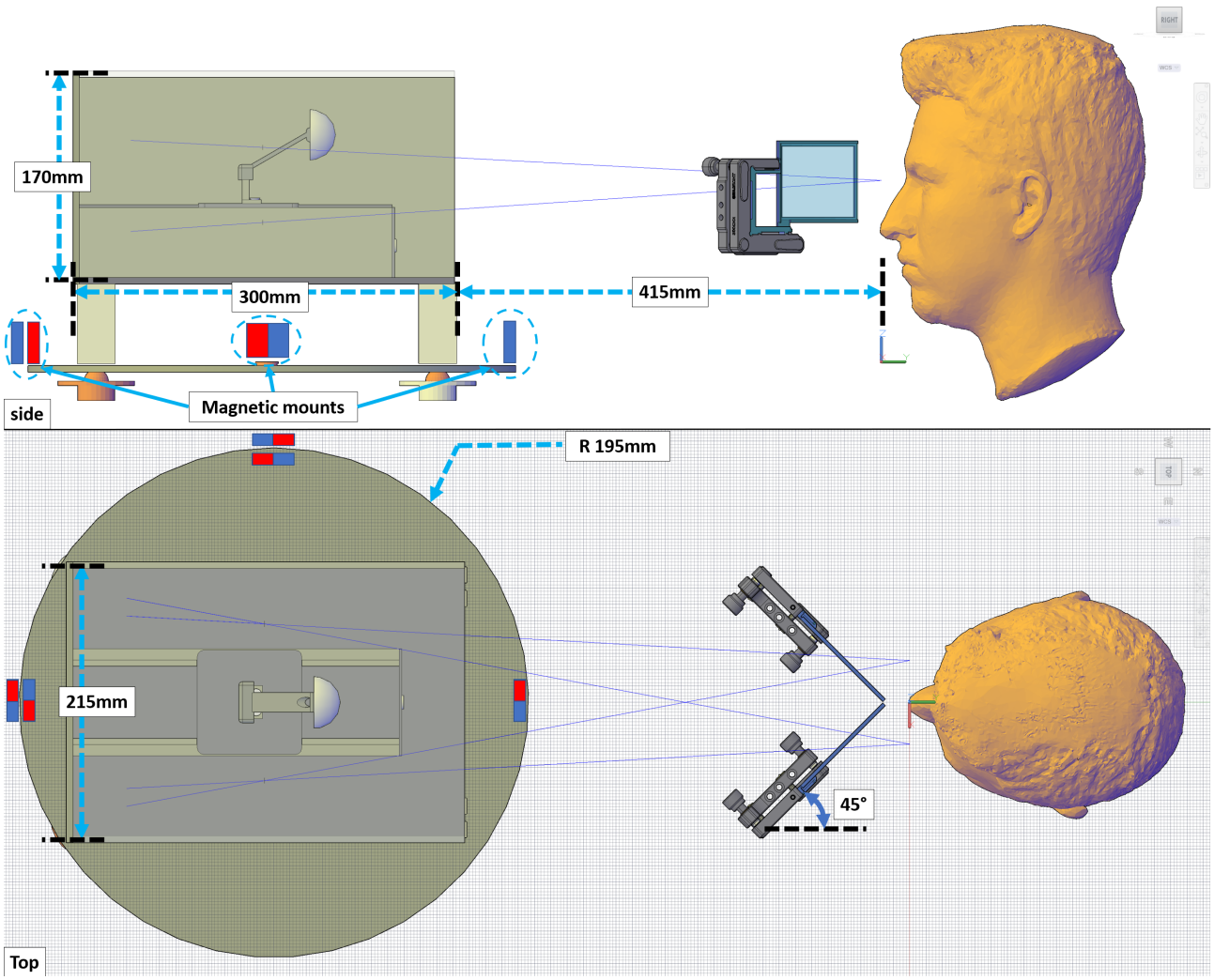

Figure 5: The CAD drawing of the real-scene box and view port.

1.3 The real-scene box

During the experiments, the real-scene box faces the observer. Figure 5 shows the top and side views when an observer is facing the real-scene box and viewing the object inside through the (70R/30T) beam-splitters. The real object is placed on a linear gantry (Oozenest, 250mm C-Beam Linear Actuator). The real-scene box is placed on a circular platform with rotating ball transfer units. Custom magnetic mounts fix the platform either towards the observer or towards the camera.

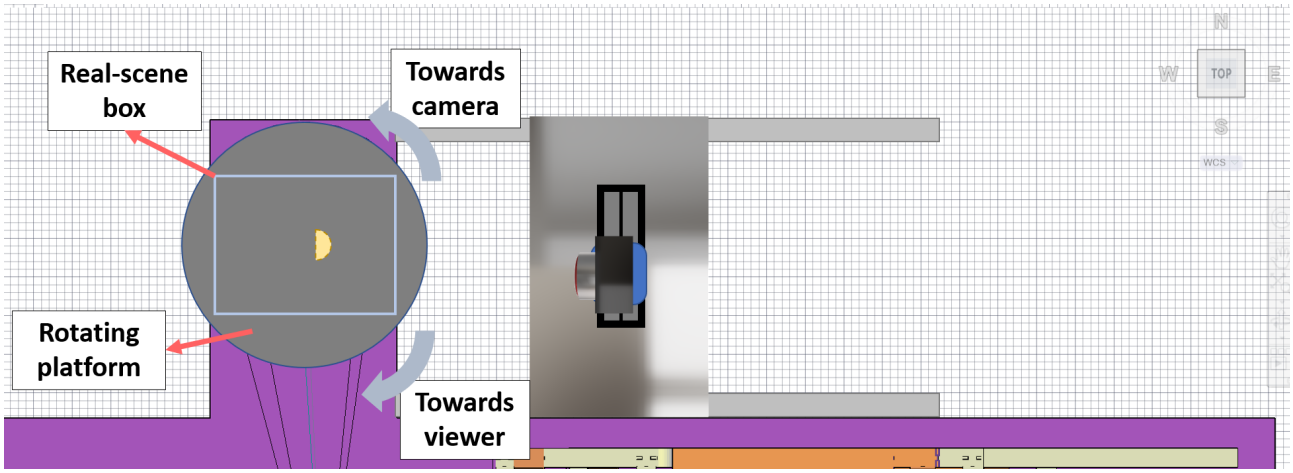

Figure 6: The CAD drawing of the camera gantry in the system.

1.4 The camera gantry

The camera-capture rig has a motorized camera gantry (shown in Fig. 6) which is capable of acquiring horizontal HDR light field data of a real scene by capturing multiple perspectives of the scene by linearly moving a DSLR camera. For capturing process, we turn the real-scene box towards the camera gantry. The capturing process is an offline step, which takes up to 1 h. The camera gantry is shown in Fig. 2 together with the real-scene box. (Note that the camera gantry is on the right hand side in the actual setup. We drew a top view of the DSLR camera, camera gantry, and the platform that we can move towards and away from the real-scene box. This reflects the actual setup.) The baseline of the camera is approximately 82.3 mm which covers a field of view much larger than the 3D objects we used.

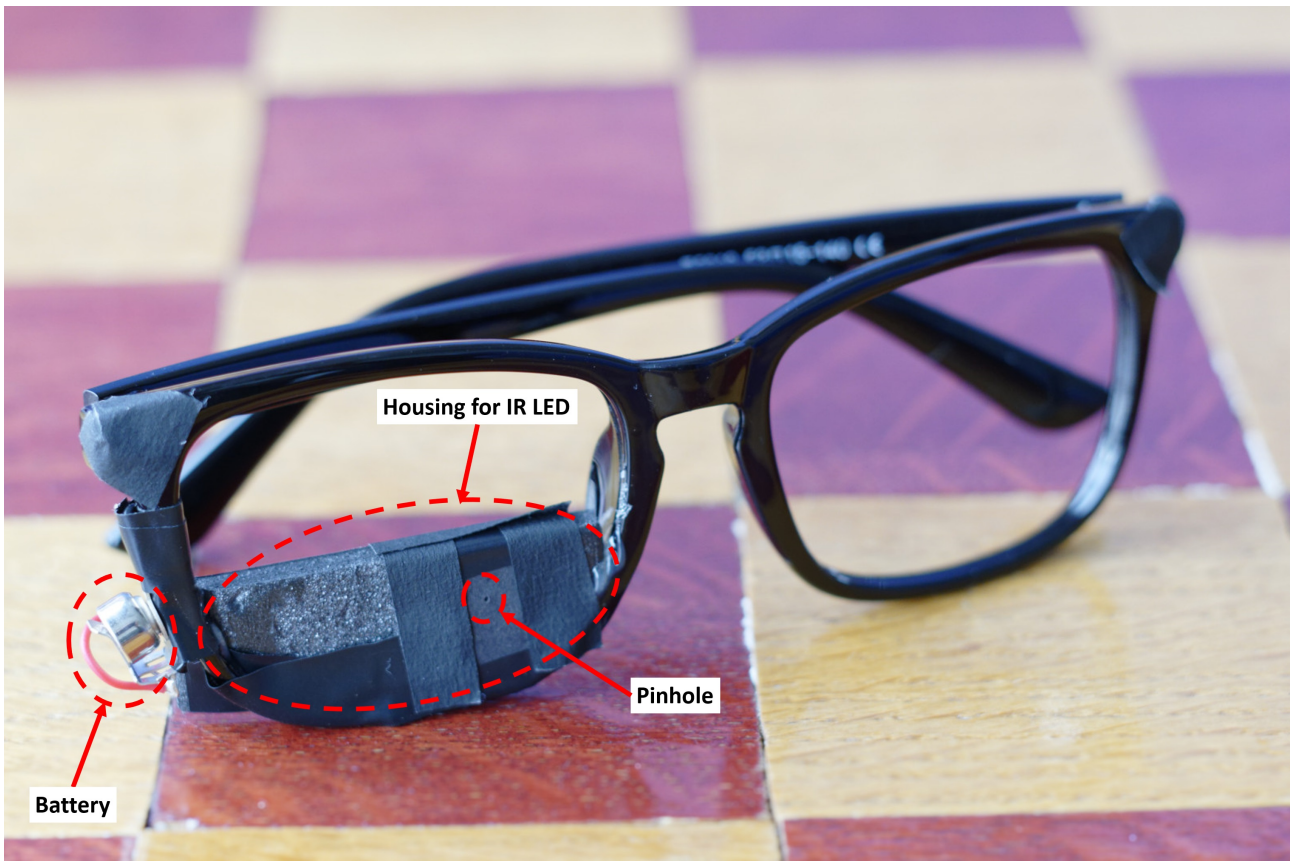

Figure 7: The glasses for head tracking.

1.5 The head tracker

The system includes a high-frame rate computer vision camera (iDS UI-3140CP), facing the eye of the participant (see Section 3.3 of the main paper). In Fig. 2, the location of this the computer vision camera is shown. The camera is meant to track the position of the head (but not gaze direction) in real-time so that it can be used to update the position of the virtual camera when the observer is moving their head. In the current experiment, the head tracking was only used to eliminate trials with unintentional head movements. The head tracker was used in combination with a customised eye-glasses frame, which included a battery operated IR LED (OSLON black series, SFH 4718A, 850 nm) emitting through a pinhole (shown in Fig. 7).

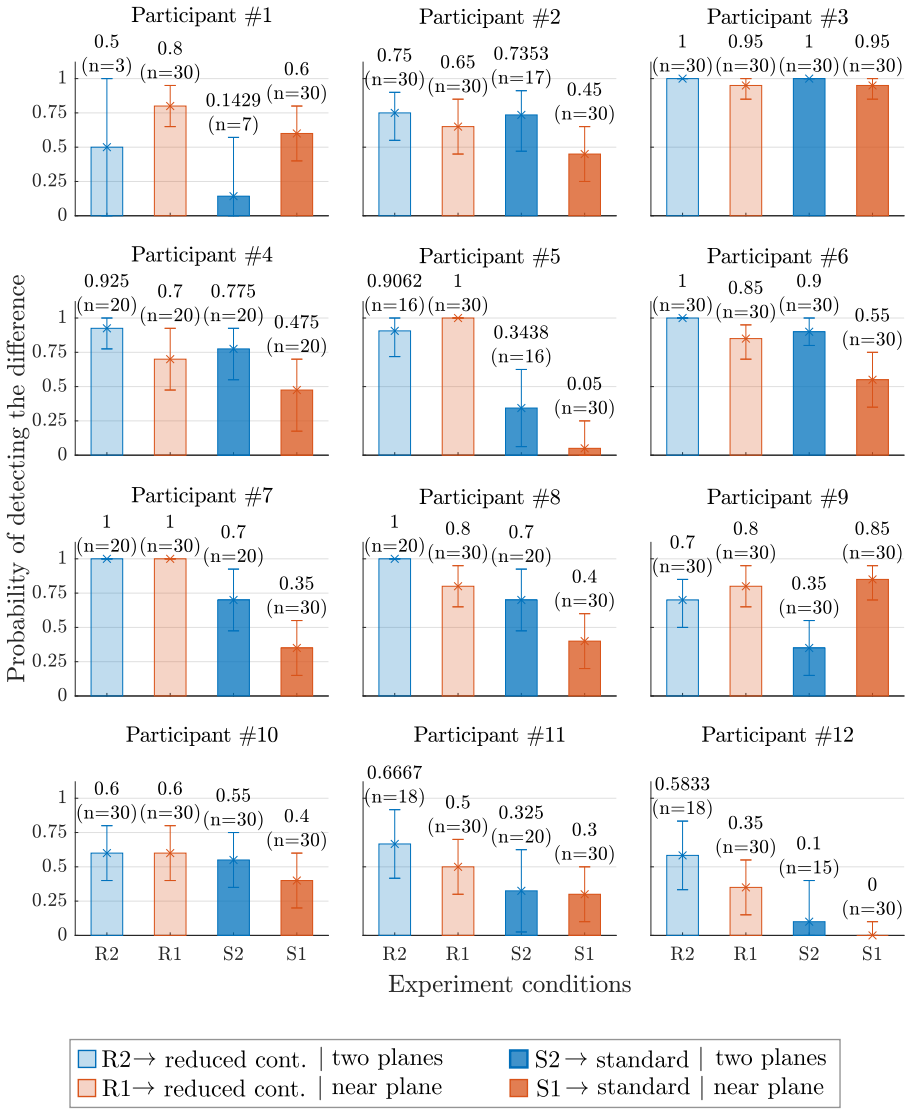

Figure 8: The probability of detecting for each condition (compensated for the guess rate) and each participant.

2 Individual performance for the visual Turing test

In this section, we present the results of each participant for our visual-Turing-test experiment (Fig. 8). In general, the results show large differences in detection probabilities across participants, which can be explained by the tendencies of different participants to pay attention to different aspects of the stimuli. In most cases, multi-focal rendering on both planes made it easier to perform the task compared to rendering only to the near plane. This could be attributed to incorrect defocus blur when accommodating between the two focal planes, since the magnitude of vergence-accommodation conflict was very small on the near plane for our experiment. Meanwhile, participants with relevant backgrounds (optics, computer graphics, psychophysics) exhibit a higher detection accuracy compared to naive observers.

In addition to the general trends, the results exhibit a few interesting individual cases. Participant #1 and #9 performed with a higher detection accuracy for multi-focal rendering, which is opposite to the main trend. However, participant #1 only had seven valid trials for standard two-plane rendering, resulting in large confidence intervals. During a post-experiment study, participant #9 indicated that the scene rendered on both focal planes better matched the real scene in terms of color and contrast. Participant #9 also reported not paying attention to the edges of the object where defocus artifacts can be more salient with multi-focal rendering. Another interesting individual result is from participant #3, who can spot the difference in almost all cases. Participant #3 was highly experienced with electronic displays, reported a 20/20 vision, and therefore was able to detect tiny differences in resolution and detail when comparing with a physical object.