Funded by the European Union

Holographic Vision for Immersive Tele-Robotic OperatioN

HoviTron

Call identifier: H2020-ICT-2019-3

Grant Agreement: 951989

D5.1 – Proof of Concept 1

Project funded by the European Commission within the H2020 Programme (2014-2020)			
Start date/ Duration	01 June 2020 / 24 months		
Deliverable leader	DLR		
Deliverable contributor(s)	Michael Panzirsch, Alex Kvasov		
Due date (DoA)	31/10/2020	**Submission date**	02/11/2020

Type		
R	Document, report excluding the periodic and final reports	x
DEC	Websites, patents filing, press & media actions, videos, etc.	x
ETHICS	Ethics requirement	
ORDP	Open Research Data Pilot	
Dissemination level		
PU	PUBLIC, fully open, no embargo e.g. web	x (report)
CO	CONFIDENTIAL, only for members of the consortium (including the Commission Services)	x (video)

1 Introduction

This deliverable presents the first proof-of-concept (PoC-1) of the visualization of a Virtual Reality (VR) environment in the light-field display developed by the company Creal. The light field approach in this Head Mounted Display (HMD) supports Eye Accommodation and VErgence (EAVE) which promises a more accurate and natural perception of depths in the VR environment when compared to standard stereoscopic displays. These properties promise to increase the usability for the human operator and to allow for a more accurate manipulation in the VR scenes, especially in terms of accurate depth perception in the near/close user region, where current headsets are unusable (they target immersive experiences mainly for distant objects). This is of high relevance in a broad spectrum of tele-operation tasks, reaching from the medical field over maintenance in space, to the dismantling of a bomb.

The following VR scenes were implemented in the cross-platform game engine Unity of Unity Technologies. In Section 2, the capabilities of the Creal light-field display in light-field rendering is presented based on pictures displaying a look of a real camera into the HMD.
Section 3 presents the currently foreseen level designs for the first user-study (US-1). In this PoC-1, it has to be proven that the light-field display is capable of displaying a large variety of virtual objects and textures. Therefore, the scenes are based on different textures, different object shapes and sizes and the involved objects are distributed in depth layers with varying distance to each other. Besides the level design presented in this report, a video of the light-field display with the developed scenes substantiates the proof-of-concept.

2 Light-field Display with Unity VR Scene

The Unity content was run, and a video thereof captured[1], which presents a look through one 'eye' of the light-field head mounted display (HMD). Via the change of the focal depth of the filming camera, the resulting realistic view of the user's eye is captured. The following pictures present screenshots from a video taken with the same method in Creal's light-field HMD.

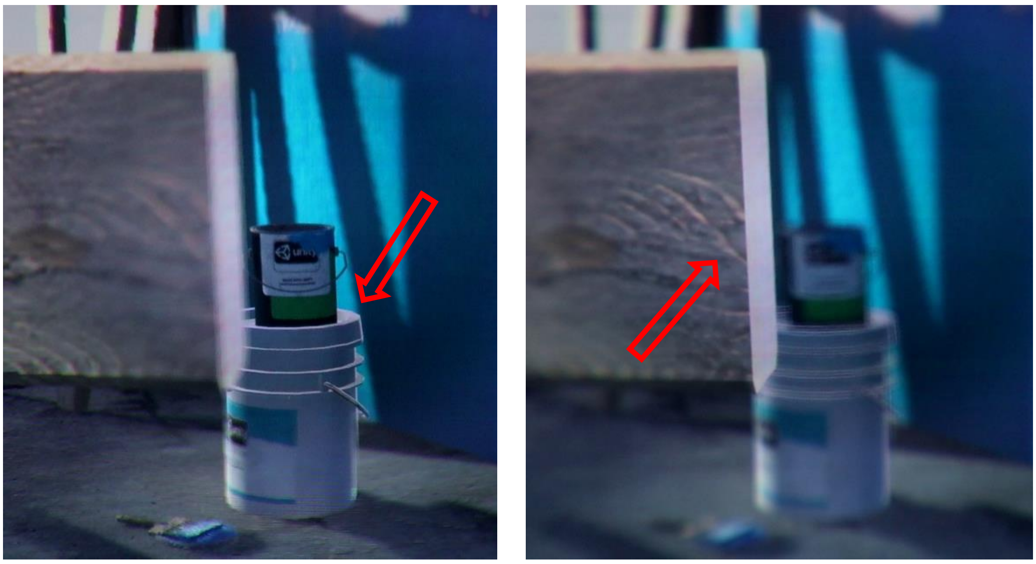


Figure 1: Scene 1

[1] The video will soon be publicly released, but not at the time of writing this deliverable, on request of some consortium members.

In Figure 1, two objects which are very distant from each other in the VR scene (0,2m to 2m distance from the virtual camera) are focused in the different views. It can be seen that if one object (e.g. the bucket) is focused, the respective other (e.g. the wood beam) is highly blurred. This focus change of the camera is analogous of the natural sensing of the human eyes when focusing on one specific object. In case of standard stereoscopic HMDs, the whole scene would be displayed sharply which reduces the perception of depth and can cause eye strain and thus negatively affects the usability.

Figure 2: Scene 1 with standard display

Figure 2 presents the same scene as Figure 1 as it would be seen through a standard HMD. If the screen of the HMD is focused by the eyes of the user, the user will see a completely sharp image (left side of Figure 2). If the focus is changed from this exact plane of the display, the whole view is blurry (right side of Figure 2). The more distant the focus is from the display plane, the blurrier the whole scene is perceived. This required constant focusing of one specific depth plane (the display plane) leads to the eye strain of standard head mounted displays and is resolved through the light field technology.

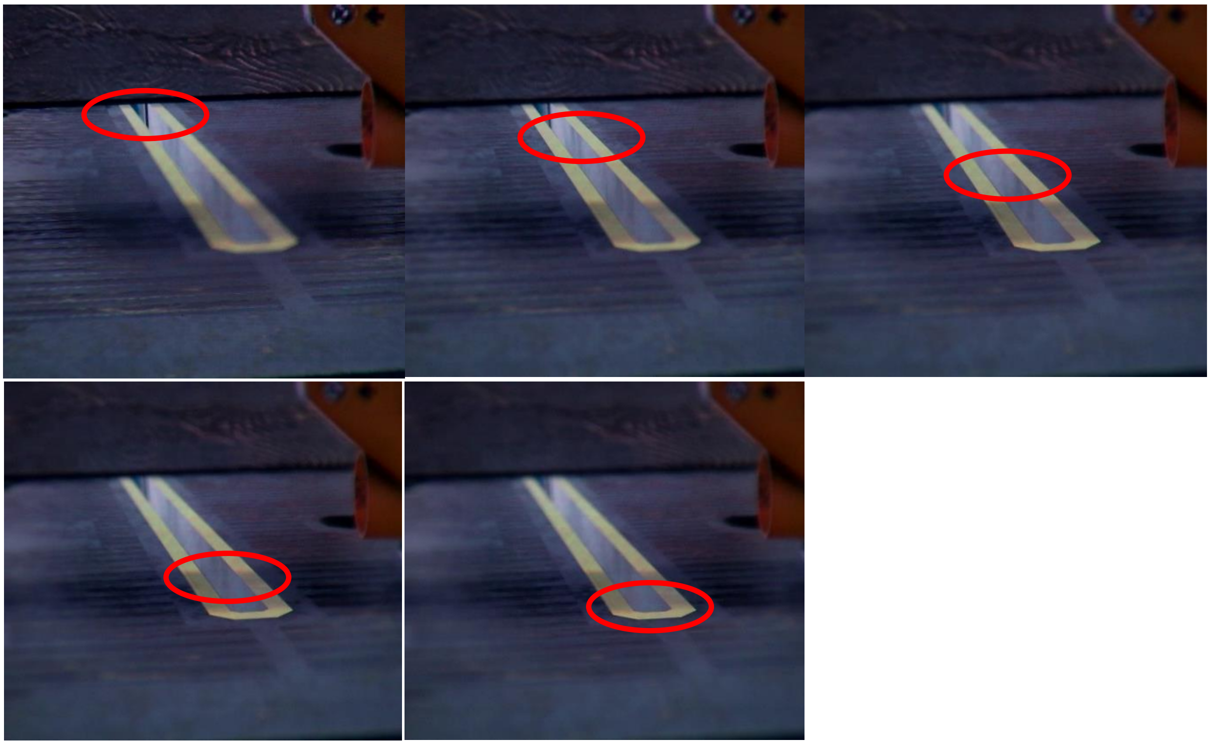


Figure 3: Scene 2

The picture sequence of Figure 3 proves the high depth resolution of the light-field display. The red ellipse marks the respective focused area in each screenshot. The visible guide rail has a length of 0.3m and is located a distance of about 0.2m from the virtual camera.

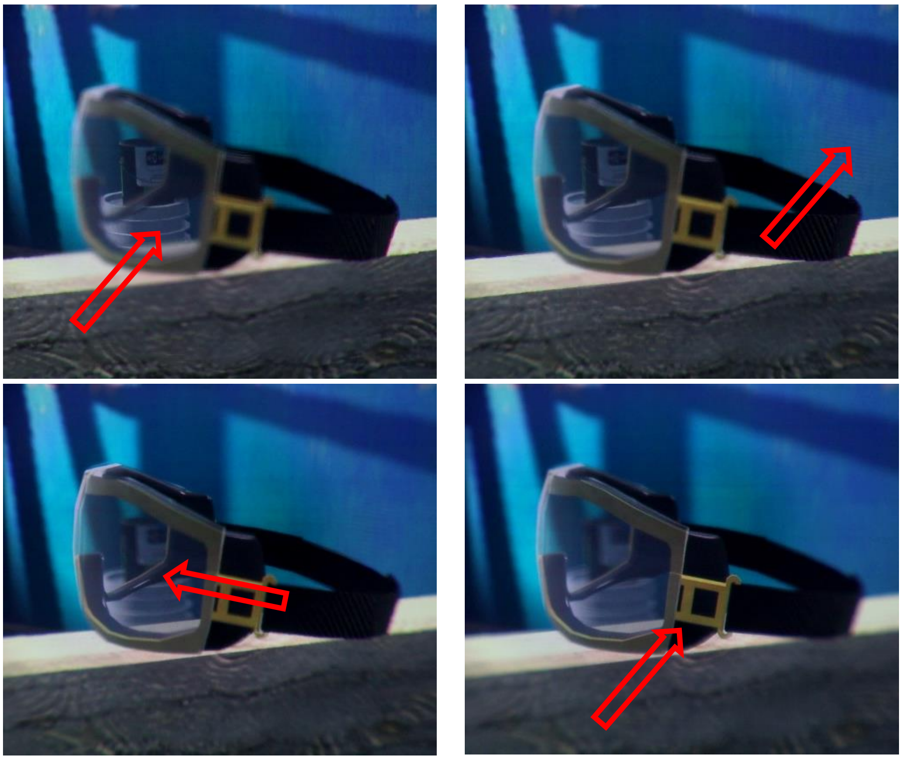


Figure 4: Scene 3

Figure 4 presents another picture sequence which underlines that the display provides a constant quality independent of the objects' texture and even when looking through transparent objects. Also, the high depth range of the visualization can be observed since the depth resolution for the safety glass is high, while also the distant bucket can be well focused.

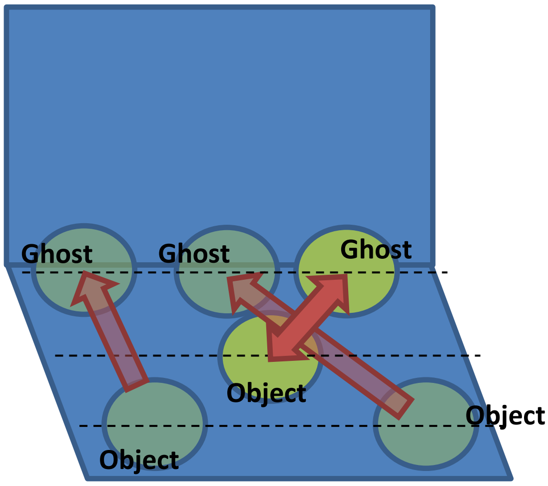
3 Level Design User Study

The scene applied in the first user study US-1 is closely related to PoC-1 also, since the capability of the display to render a large variety of textures and types of objects needs to be assured. The following three different tasks/levels were designed for the first user study.

3.1 Aiming Task

Aiming tasks are one of the main classical tasks in the performance evaluation of VR and tele-robotic systems. The following box presents the abstract concept of the task to highlight its relevant aspects. The blue planes represent a vertical and horizontal plane in the VR scene representing the background.

TRANSLATORY AIMING



Task description:

The position of the **Object** is coupled to the haptic user interface. The task is to match the position of the object with a **Ghost** object that appears at different locations in the workspace. Only translations but no rotations have to be set such that the speed of depth adaptation plays the crucial role. Perpendicular and cross motions are considered.

This abstract concept was implemented in a more realistic and motivating setup as presented in Figure 5 and Figure 6.

Figure 5: Bird view

Figure 6: Operator View

Task:
The position of one apple which is controlled via the haptic input device has to be matched to different appearing ghost apple locations.

Challenge:

- Positioning accuracy
- Execution Time

Expectation:

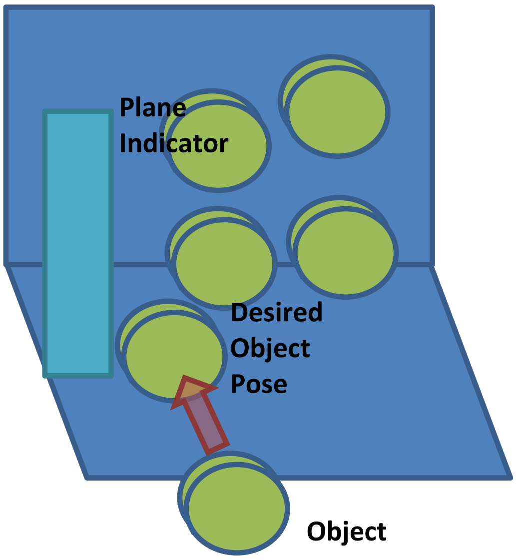
A light field…

- …provides a more convenient experience
- …eases the position match in terms of time and accuracy

Also, the expected faster perception of the depth of the goal should lead to an earlier start of motion (e.g. also higher acceleration).

3.2 Depth Perception Task

In contrast to the aiming task, the following task was designed especially according to the specific capabilities of the light-field technology.

Disc Array

Task description: The different disks of the disk array have to be approached with the **Object** in a specific order, e.g. first the depth of the most distant disk then the depth of the closest disk then the depth of the second most distant disc and then the depth of the second closest disk and so on, has to be matched.

This abstract concept was implemented in a more realistic and motivating setup as presented in Figure 7 and Figure 8.

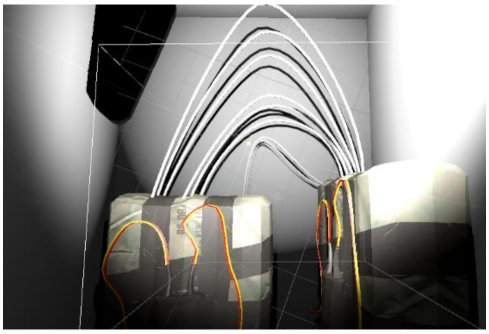


Figure 7: Bird view

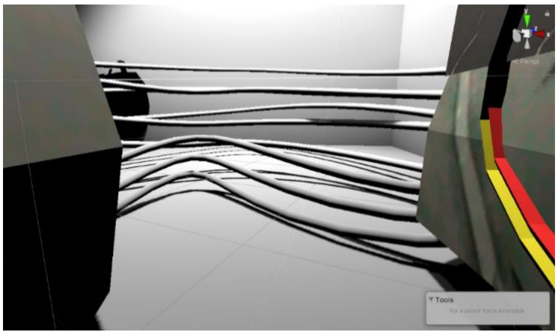


Figure 8: Operator View

Task:

In this task, a bomb has to be disarmed which represents a tele-operation procedure which could benefit from the developed light-field technology. The wires have to be cut in a specific depth order. Therefore, the wires have to be touched for a specific time with e.g. a virtual scalpel which is controlled via the haptic input device.

Challenge:

- Correct evaluation of depth
- Shadows can be added to render the task more difficult

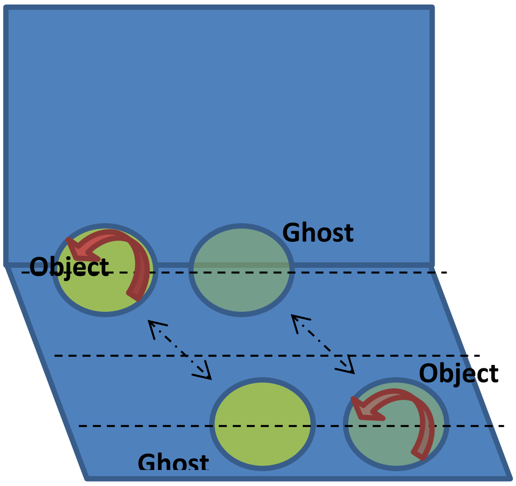
Expectation:

A light-field…

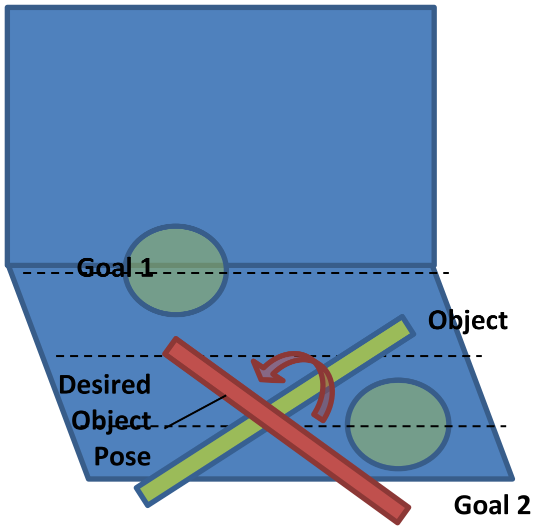
- …allows the fast and correct evaluation of distances

3.3 Depth Variation Task

This task provokes a frequent change of view depth through implementation of a manual handling in the background of the scene. This subtask is combined with a second challenge which is the positioning of two points of interests in different depth planes.

Manual handling

Task description: The orientation of an **Object** has to be matched with the orientation of a **Ghost** object. The **Ghost** object is presented in a different depth plane similar to a manual handling for assembly in real life.

Two-layer depth reference



Task description:

The endings of a long **Object** have to be positioned in sphere-like **Goal** positions. The **Goal** positions are located in different depth planes.

The implementation of these subtasks is visualized in Figure 9 and Figure 10. The red boxes mark the manual-like assembly (top-view) in the background and the table on which the parts to assemble are located.

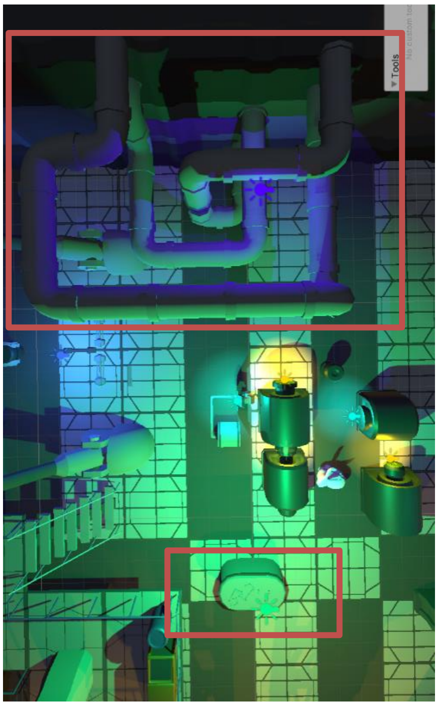


Figure 9: Bird view

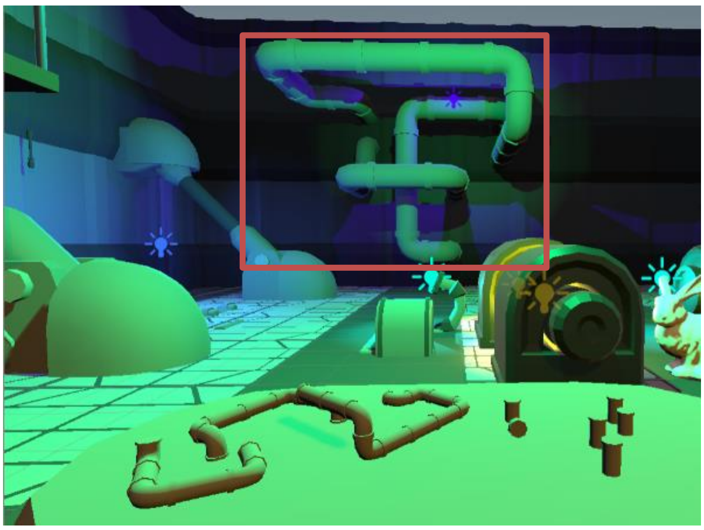


Figure 10: Operator View

Task:
The pipes have to be assembled analogous to the structure in the background which presents the top view on the desired structure.

Challenge:
The iterative focusing in the depth of the table and the background to analyse the correct pipe assembly is more tedious due to eye strain without the natural light-field technology.

Expectation:
A light field …

- …provides a more convenient experience when switching focus between background and table
- …helps to position the pipe end at the two fixed sockets located in different depth planes

Outlook

The PoC-1, including the level designs presented in Section 2, builds the baseline for the user-study US-1. The detailed design and results of the user-study are content for deliverable D5.2 which is due end of March 2021. First documents on administrative aspects of US-1 will be discussed in D7.1 and D7.2.