

MPEG Immersive Video tools for Light-Field Head Mounted Displays

Daniele Bonatto

*Laboratories of Image Synthesis and Analysis
Université Libre de Bruxelles
Brussels, Belgium
daniele.bonatto@ulb.be*

Sarah Fachada

*Laboratories of Image Synthesis and Analysis
Université Libre de Bruxelles
Brussels, Belgium
sarah.fernandes.pinto.fachada@ulb.be*

Grégoire Hirt

*CREAL
Lausanne, Switzerland
gregoire.hirt@creal.com*

Alexander Kvasov

*CREAL
Lausanne, Switzerland
alex.kvasov@creal.com*

Gauthier Lafruit

*Laboratories of Image Synthesis and Analysis
Université Libre de Bruxelles
Brussels, Belgium
gauthier.lafruit@ulb.be*

Abstract—Light-Field displays project hundreds of micro-parallax views for users to perceive 3D without wearing glasses. It results in gigantic bandwidth requirements if all views would be transmitted, even using conventional video compression per view. MPEG Immersive Video (MIV) follows a smarter strategy by transmitting only key images and some metadata to synthesize all the missing views. We developed (and will demonstrate) a real-time Depth Image Based Rendering software that follows this approach for synthesizing all Light-Field micro-parallax views from a couple of RGBD input views.

Index Terms—Demo, View Synthesis, DIBR, Holography

I. INTRODUCTION

This paper presents a demonstration centered on the topic of view synthesis that is complementary to the tutorial session “The MPEG Immersive Video coding standard” to be held at the Visual Communication and Image Processing (VCIP) conference 2021. Our approach [1], [2], we will demonstrate at VCIP for Light-Field VR, is based on Depth Image Based Rendering (DIBR): the pixels from input views are first projected to 3D space using their depth maps, and afterwards reprojected to any virtual view one wishes to synthesize.

II. HOLOGRAPHIC STEREOGRAMS

A holographic stereogram is a Light-Field display for static objects, where hundreds of micro-parallax images are rendered in adjacent directions from laser-engraved interference fringes. In [3] we have shown that MIV’s Depth Estimation Reference Software (DERS) and Reference View Synthesizer (RVS) allow to create these micro-parallax images, starting from only four RGBD input images. Though the depth image (D) itself needs a couple of more input images to reach high quality standards with depth estimation (DERS), the total number of RGB input images required to synthesize all viewpoints of the holographic stereogram with DERS and RVS is limited to maximum a dozen. The procedure to make holographic stereograms using our software can be found in [4].

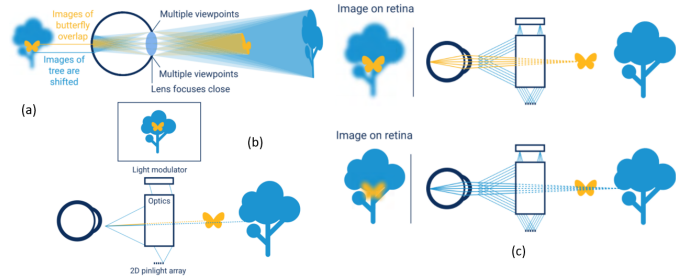


Fig. 1. Light-Field rendering (a) with time-multiplexed light modulator (b), yielding focus points at various depths (c).

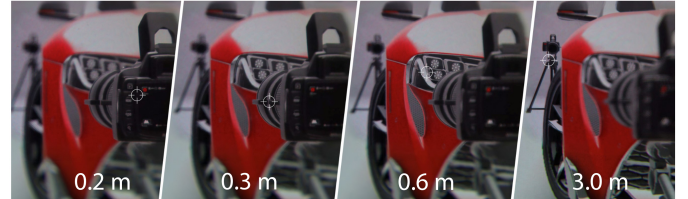


Fig. 2. Creal’s rendered Light-Field on a HMD at various focus points, cf. the cross-haired white circle in each image.

III. VIEW SYNTHESIS IN STEREOSCOPIC HMD

Strengthened by the holographic stereogram experience of previous section, we decided to port our solution to a stereoscopic Head Mounted Device (HMD) for Virtual Reality (VR). Though only two virtual views following the user’s pose must be rendered at any time, the high frame rates used in VR (60-120 fps) impose stringent real-time constraints. Optimizing an in-house OpenGL implementation of RVS - called RaViS - we reach real-time performances, generating two synthesized views at 60 to 90 fps [2].

IV. VIEW SYNTHESIS IN LIGHT-FIELD HMD

In the context of the European project HoviTron, Grant Agreement 951989, H2020-ICT-2019-3, we were faced with the challenge of projecting a Light-Field per eye in the HMD

for overcoming the eye accommodation and vergence conflict, which causes cybersickness. Instead of projecting a single image per eye, various images with micro-parallax are projected into each eye, as shown in Fig. 1(a). The images presented on a light modulator (top of Fig. 1(b)) are slightly shifted with a time-multiplexed pinlight array (bottom of Fig. 1(b)), virtually creating focal points at various depths, cf. Fig. 1(c). This mimics holographic wavefronts that are orthogonal to the light rays entering the eye, recovering the ability to the user to focus on any point in space at will, without any eye tracking system that would have to blur image regions to simulate a Bokeh effect. Preliminary user tests [5] conducted at Deutsches Zentrum für Luft und Raumfahrt (DLR), Germany, confirm these findings, also shown in Fig. 2 for a synthetic, yet photo-realistic scene, rendered in Creal's Light-Field HMD, photographed at various focus points, as if a user would look at these targets through the HMD.

However, many more images must be rendered instantaneously in such Light-Field HMD, which cannot be reached with 3D graphics software, readily available on the market. Also RaViS reaches data rate bottlenecks in the rendering pipeline, which could only be overcome by exploiting temporal image redundancies and advanced caching techniques, all packaged into Creal's Spatio-Temporally Amortized Light-Field (STALF) approach.

REFERENCES

- [1] S. Fachada, B. Kroon, D. Bonatto, B. Sonneveldt, and G. Lafruit, "Reference View Synthesizer (RVS) 2.0 manual, MPEG2018/N17759," *ISO/IEC JTC1/SC29/WG11*, Jul. 2018.
- [2] D. Bonatto, S. Fachada, and G. Lafruit, "RaViS: Real-time accelerated View Synthesizer for immersive video 6DoF VR," *Society for Imaging Science and Technology (IS&T) - Electronic Imaging*, Jan. 2020.
- [3] S. Fachada, D. Bonatto, and G. Lafruit, "High Quality Holographic Stereograms Generation using four RGBD Images," *Applied Optics*, Oct. 2020.
- [4] S. Fachada and G. Lafruit, "Computer Generated Holography with Depth-based View Synthesis," in *Imaging and Applied Optics Congress*, Jun. 2020.
- [5] M. Panzirsch, B. Weber, and M. Lingenauber, "HoviTron - D5.2 - User Study 1 (US-1), European project HoviTron, H2020-ICT-2019-3, Grant Agreement 951989," Jul. 2021.