A Mixed Reality Generator for Real-World Environments in Real Time

Devrim Akca[1], Çağın Torkut[2], Gerhard Kemper[3], Armin W. Grün[4]

[1] Faculty of Engineering and Natural Sciences, Işık Üniversitesi – devrim.akca@isikun.edu.tr
[2] RedHorizon Technology, Inc., – torkut@redhorizon.net
[3] GGS GmbH,Speyer / Germany – kemper@ggs-speyer.de
[4] 4DiXplorer AG – armin.gruen@geod.baug.ethz.ch

KEYWORDS: Real-time stereoscopic visualization, 3D stereoscopic mode, Extended Mixed Reality, Camera-calibration, UAV, Monitoring critical infrastructure

ABSTRACT
By integrating computer vision, photogrammetry, UAV and Extended Reality (XR) solutions, the proposed innovative Mixed-Reality (MR) photogrammetry system enables real-time 3D visualization, interaction and measurement of real-world environments. Eliminating the need for physical presence, the system enhances safety, efficiency and accuracy in tasks like assessing structural integrity, tracking construction progress, and observing environmental changes over time. At the core of the system is a UAV equipped with a stereo camera rig and onboard processing capabilities. Operated on-site by an operator, the UAV captures high-resolution stereo imagery processed in real time through a centralized Rest API running on cloud infrastructure. Experts and engineers located anywhere in the world can connect to the system using VR headsets or a web-based application, gaining immersive access to a 3D stereoscopic view with full photogrammetric measurement functionality. The system supports multi-user collaboration, enabling synchronized analysis and data sharing across different locations. This seamless integration of hard- and software components represents a significant advancement in real-time stereoscopic visualization. Key features include a 2-camera stereo rig mounted on the UAV, mixed-reality visualization software for VR, and a web-based interface. Combining real-time data acquisition, processing, and visualization, the system provides unparalleled insights, surpassing the limitations of traditional on-site inspections. Designed with focus on cost-effectiveness and safety, this innovative approach is particularly valuable for inspecting large-scale, critical structures in challenging or hazardous environments. The proposed system offers a transformative tool for remote monitoring, ensuring timely and informed decision-making in diverse applications.

1. BACKGROUND

Protection of critical infrastructures (such as power grids, transport networks, information and communication systems) is vital for the security of countries and the well-being of their citizens. Infrastructure elements such as bridges, viaducts, dams, towers, etc., are exposed to intensive loads, unsteady weather conditions especially due to climate change, material fatigue and aging process. All these factors cause deformations & deteriorations on the construction parts (e.g. Figure 1) and can cause a high social and economic impact. The developed solution shows a monitoring & inspection system with a stereo camera equipped UAV for on-demand inspection.

Figure 1. monitoring critical infrastructure under real-time conditions with remotely connected specialists is the goal of this novel technology

Due to economy & safety reasons, inspection work of large civil and critical structures is envisaged in such a way that the expert team observes the object of interest using VR headsets in-office while an operator runs a stereo camera equipped UAV on-site. Core of this inspection system is an MR based photogrammetry software which allows the remote users to access, to visualize and to measure the stereoscopic model simultaneously.

2. INTRODUCTION

Mixed reality (MR) blends real and digital worlds offering innovative solutions across various fields such as from architecture & construction to movie industries. It enhances and augments human perception and interaction with the world. Looking at the number and value of the venture capital (VC) investments, which is a valuable metric for commercial innovation, the growth of activity in augmented, virtual and mixed reality (AR/VR/MR) has been consolidated in the 2012-2015 period before their parabolic run (The Economist, 2025).

Basic data source of the MR systems is the image which can be acquired by different sensors. We foresee a future trend where image acquisition sensors and processing devices become increasingly integrated. This next era, referred to as real-time photogrammetry, involves on-the-fly processing performed either on PCs located near the sensors on-board the platform or in the cloud. As computational power and data storage shift away from traditional office setups, workstations are being replaced by simple terminals, monitors, VR headsets, and MR glasses.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-1005-2026

This paper introduces one of the first prototypes in this emerging field, focusing on the development of a live-streaming, multi-user MR photogrammetry system. The system incorporates real time 3D stereoscopic visualization and utilizes a stereo camera-equipped Unmanned Aerial Vehicle (UAV) as its sensor platform.

3. UAV & MR

Critical Infrastructure encompasses various facilities such as transportation systems, electricity networks, telecommunications, and water supply structures. These elements face pressures from natural forces and human activities, necessitating consistent monitoring. Their construction involves significant complexity and cost, and unexpected damages can lead to substantial economic losses and potentially catastrophic consequences. Therefore, continuous monitoring & surveying of infrastructure objects—such as bridges, viaducts, dams, and towers—are essential. These surveys play a vital role in ensuring public safety, maintaining structural integrity, and guiding necessary rehabilitation efforts.

UAVs are among the most affordable and widely used platforms for collecting image data for such monitoring tasks. They have transformed both industry practices and academic research trends due to their flexibility, portability, and affordability (Colomina and Molina. 2014, Noor, Abdullah and Hashim. 2018, Nex, Armenakis, Cramer, Cucci, Gerke, Honkavaara, Kukko, Persello, and Skaloud. 2022). The terms UAV, Unmanned Aircraft System (UAS), and Remotely Piloted Aircraft System (RPAS) are often used interchangeably, leading to occasional confusion, as discussed by (Granshaw. 2018). UAV photogrammetry holds significant potential for achieving high-accuracy results, often surpassing those of conventional digital airborne photogrammetry (Forlani, Dall’Asta, Diotri, Cella, Roncella and Santise. 2018).

UAVs have been utilized for a wide range of monitoring and visualization applications, including the open-pit mine monitoring, energy infrastructure damage assessment, power line inspection, crack detection, bridge inspection and visualization of historic structures (Cao, Zhang, Zhang, Yin, Man 2023, Bowman, Yang, Thomas, Kirk, Duncan, Hughes and Meade. 2023. Li, Zhu, Guo, Wang, Jiang, Wang and Liao. 2022. Meng, Gao, Zhou, He and Djerrad. 2022. Zhong, Peng, Yan, Shen, and Zhai. 2018.). Images are typically collected in a strip or block formation with specified forward and side overlap ratios. These images are then post-processed in subsequent steps to generate detailed 3D information.

Our research focuses on the real time MR generation for real world environments and objects. Against to this conventional (off-line) methodology, our approach turns the UAVs into a platform to obtain real time capability. It integrates a stereo camera rig as the image acquisition system, mounted on an UAV (or "drone"), MR technology, photogrammetry and computer vision (Figure 2). The stereo base length was designed following photogrammetric base-to-height ratio principles, enabling real time 3D stereoscopic visualization and measurement provided that the left and right images are fully synchronized. The “real time” capability is made possible through the combined use of an onboard PC, a 4G/5G modem, cloud computing technologies and the VR headsets.

A real time concept was presented by Xiao et.al. (Xiao, Nouri, Hegland, Garcia, and Lanman. 2022) which utilizing a custom-built VR headset with a stereo RGB camera rig. Their method delivers real-time, perspective-correct stereoscopic views from the stereo cameras fixed on the VR headset. Real-Time omnidirectional stereo rendering using the panoramic images is done by Marrinan and Papka (Marrinan and Papka. 2021).

Figure 2. UAV as a MR generator equipped with the stereo cameras over a 1m base length

The aim of our study is to generate 3D stereoscopic views by processing stereo camera images onboard a UAV and transmitting them via the UAV's onboard PC, through a 4G/5G modem and cloud infrastructure, to the final destination—the VR headsets of the users. In our scenario, the UAV system is operated on-site by a pilot, in real time, the users conduct visualizations, measurements & inspections remotely using VR headsets. This is achieved by tight integration of the hardware and software components. This approach is unique as it seamlessly integrates data acquisition and photogrammetric image processing into a unified, real-time workflow. The photogrammetric outputs are interactively and simultaneously shared among multiple users, regardless of their geographical locations. Moreover, it provides an exceptionally rich visualization experience, surpassing the level of detail achievable by physically being at the site.

The second section provides an overview of the system, detailing its individual components in terms of both software and hardware. The third section presents the first results. Finally, the fourth section summarizes the conclusions.

4. MR GENERATOR

4.1 Components overview

The system primarily comprises a UAV equipped with stereo cameras and an MR software suite, specifically designed for real time inspection and visualization (Figure 3). The hardware and software components are seamlessly integrated into a unified pipeline, enabling real time image acquisition and photogrammetric processing. The system includes the following items:

- A UAV equipped with stereo cameras and onboard PC which runs middleware software,
- A MR photogrammetry application running on VR headsets,

- A backend system and web application hosted on cloud servers.

While asynchronous monoscope cameras (for both photo and video) are widely used on UAVs, the application of simultaneous stereoscopic camera systems in the geospatial market is both novel and uncommon. This is a new system since all the relevant technologies have never reached a productive level.

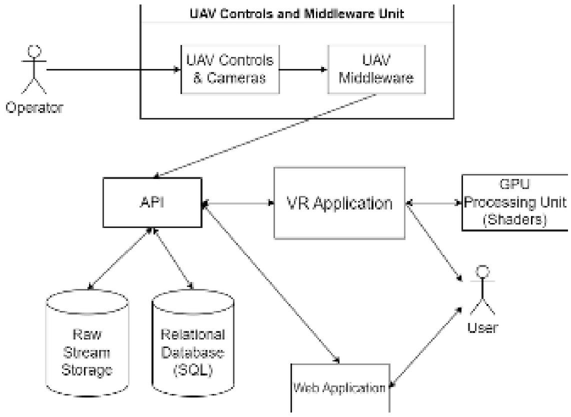

Figure 3. System architecture. Multiple users from different geographical locations can simultaneously access data via the system.

4.2 UAV

The GGS AeroSpector (AS) 800 UAV serves as the aerial platform, featuring a foldable carbon quadcopter frame capable of carrying payloads of up to 2 kg. It is a fully assembled, ready-to-fly (RTF) drone equipped with a Herelink remote control system. The AS-800 has a motor-to-motor diameter of 800 mm. Without a payload, the drone achieves a flight time of approximately 70 minutes. With a 500 g payload, the flight duration is reduced to about 52 minutes, and with a 1,000 g payload, it decreases further to 45 minutes. The drone's base weight, including all electronics, is just 1,800 g. When equipped with a 1,000 g payload and a 17,000 mAh GensAce battery, the total take-off weight remains under 4.5 kg (Figure 4).

Figure 4. The UAV and the stereo cameras

The system utilizes two 20 MP Alvium cameras (Alvium 1800 U-2040c models), featuring an USB 3.1 interface, a C-Mount lens connector, and a CMOS sensor with 4,512 x 4,512 pixels at a pixel size of 2.74 µm (Figure 5). Both cameras are paired with 16 mm high-resolution lenses from the same manufacturer. The cameras are mounted on either end of a 1-meter carbon tray (stereo rig) and tilted approximately 5 degrees to achieve a convergent imaging geometry.

Figure 5. One of the high-resolution stereo cameras mounted on the rig

They are connected to an onboard PC (Latte Panda 3D), which manages image pre-processing and data transfer. Powered via a USB-3 interface with Power over Ethernet (PoE), the cameras are controlled through the Vimba SDK software development kit. A separate port connects a camera trigger interface to the UAV controller. Additionally, GPS-RTK geotags and inertial measurement unit (IMU) data is streamed in real time alongside the images. In order to reduce vibrations, a special concept of dampers have been integrated in a way, that no additional lever-arm resonances appear. This was a rather long way with several Lab analytics and field tests to prevent good image quality without reducing flight-safety and autonomous flight control based on GNSS-INS and accelerometers.

4.3 MR application

Developed using the Unity 3D engine, the MR application serves as the interface through which users can engage with the 3D stereoscopic models generated by the system. It seamlessly integrates with the backend architecture and hardware components, providing a highly detailed and interactive experience for users across various industries. The application leverages OpenXR for cross-platform compatibility and has been tested and verified on Oculus Quest 3 to ensure a seamless and user-friendly experience. The MR application offers the following core functionalities:

- Real time visualization: Utilizing live stereo image streams from the UAV-mounted cameras, the application generates 3D stereoscopic models that can be visualized in real time using VR headsets. This enables users to interact with detailed representations of real-world environments without requiring physical presence.

- Interactive Tools: The application includes tools for labeling, annotating, measuring, and vectorizing objects within the 3D models. These tools facilitate detailed analysis, enabling precise inspections and decision-making.

- Multi-User Collaboration: The MR application supports simultaneous access by multiple users from different geographical locations. By leveraging cloud computing and a centralized Rest API, it ensures synchronized visualization and interaction for collaborative workflows.

- Photogrammetric Measurement Integration: The MR application integrates advanced photogrammetric functionalities, allowing users to perform accurate measurements within the 3D models.

To achieve its advanced capabilities, the MR application leverages a combination of computer vision and graphics algorithms, including:

(1) Real time stereo image normalization,
(2) Pyramid-based block matching with normalized cross-correlation,
(3) Precise Spatial Intersection for 3D coordinate computation.

The MR application features a native 3D GUI designed specifically for VR headsets. This interface provides intuitive controls for navigating the 3D environment, interacting with objects, and accessing analytical tools. The design emphasizes user comfort and accessibility, ensuring an engaging experience for users of varying expertise levels. Unlike traditional systems that rely on cartographic abstractions such as 2D maps or static 3D models, the MR application prioritizes augmented stereoscopic models. These models are stored and retrieved from the database in their enhanced format, offering unparalleled detail and contextual awareness. This approach significantly improves the user's ability to assess and interact with real-world environments in virtual space.

The modular architecture of the MR application ensures scalability and adaptability for future advancements. By integrating additional sensors, enhancing data processing algorithms, or expanding the scope of supported industries, the application can evolve to meet emerging needs. Furthermore, its compatibility with web-based interfaces extends its accessibility to users without VR hardware, providing measurement and annotation functionalities on 2D images through a browser-based platform.

4.4 Web application

The web application serves as an essential component of the system, providing access to key functionalities for users who may not have access to VR headsets. Designed with simplicity and efficiency in mind, the web platform ensures compatibility across a variety of devices and browsers, enabling broad accessibility. While the MR application focuses on immersive 3D visualization, the web application complements it by offering lightweight, browser-based functionality. This dual approach ensures that the system remains accessible to a diverse range of users, accommodating both high-end VR setups and standard web-based workflows.

5. APPLICATION

5.1 Powder Keg Tower in Wissembourg

A cultural heritage object in Wissembourg (France) was used in the application. The tower, known as the "Tour de la Poudrière" (or "Powder Keg Tower" in English), built in the 14th century, marks the northeast corner of the historic city of Wissembourg, and is part of the city wall. This tall, square tower, featuring capped levels from the 13th and 14th centuries, was open at the rear and served to control the northeast corner of the medieval curtain wall. It continued to function as an observation post during the Second World War as part of the "Magino Line". In the 16th century, it was preceded by a bastion designed for artillery, reinforcing its strategic importance. The tower remained a key feature of the updated fortifications of the 18th century. The rampart ditch was equipped with a lock system, allowing it to be flooded when necessary. At the base of the rampart, a vaulted underground passageway leads to the Rue de la Laine.

5.2 Stereo camera calibration

The stereo cameras onboard the UAV are designed to provide 3D stereoscopic vision, and prior calibration is essential. A calibration flight for photogrammetric camera calibration is conducted before commencing each new project. For the Powder Keg Tower project, the calibration flight was done on October 21, 2024 before the real time image acquisition of the object Figure 6). Two different camera viewing modes were utilized to strengthen the network geometry. A total of 20 stereo image pairs were captured in the nadir view and another 20 pairs in the oblique view to fully cover the test-field (Figure 7).

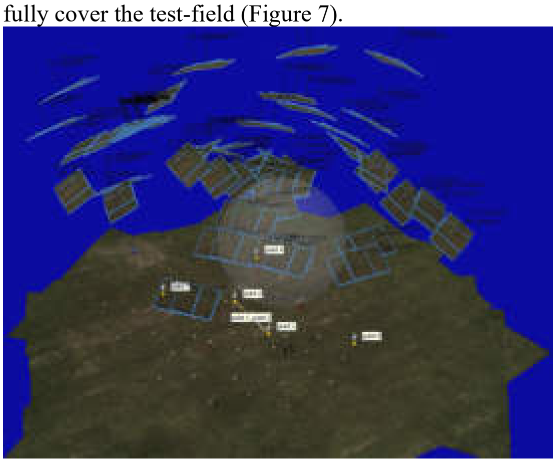

Figure 6. Image alignment of the 3D test-field using targets also on tripods for a better focal length calibration

Figure 7. The left and the right camera images, precisely synchronized, captured in the oblique view

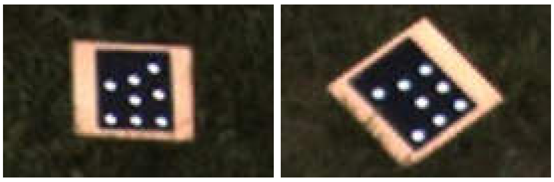

Figure 8: Enlarged views of the coded targets seen in Figure 7.

The average flight height was approximately 12 meters. Coded targets from the Australis photogrammetry software were utilized for automatic point measurement (Figure 8). The self-calibrating bundle adjustment estimated the accuracy of the 3D point coordinates as ± 0.13 mm (or 1:109'900). The interior orientation parameters and lens distortion parameters for both the left and right cameras were simultaneously computed as shown in Figure 9. Additionally, the dependent relative orientation parameters between the left and right cameras were determined.

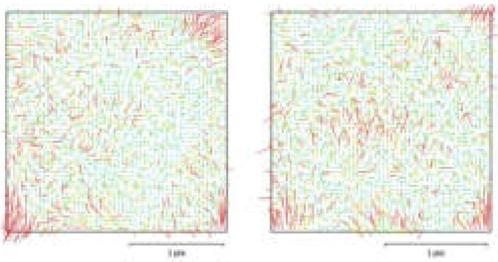

Figure 9: Left and right camera calibration residuals

5.3 Image normalization for stereoscopic visualization and tools for measurements

Image normalization is essential for achieving 3D stereoscopic vision (Figure 10). This process involves resampling the left and right images into a state where the y-parallax is eliminated. The two normalized images share common rotation angles (omega, phi, and kappa). In these images, the epipolar lines in both the left and right images align to have the same y-coordinates, effectively eliminating the y-parallax.

Figure 10. Stereoscopic view including measuring tools of the Powder Keg Tower

Besides that, the reconstruction of digital twin as shown in Figure 11 can be post-processed as well for additional documentation

Figure 11. Postprocessed to a digital twin after stereoscopic inspection using the streamed image data

6. CONCLUSIONS

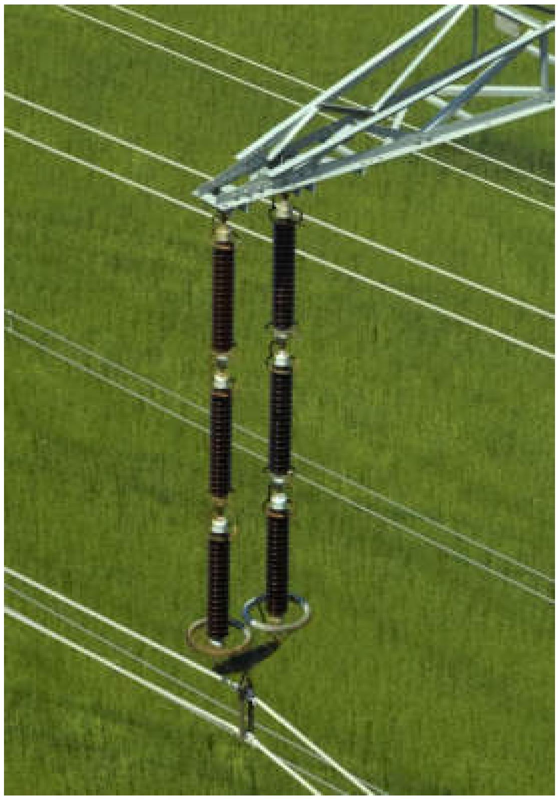

Figure 12. Inspection of powerlines can be done with this system in real-time without post-processing

Inspection of infrastructure elements, as shown in Figure 12 as a powerline example, are often outsourced to specialized engineering companies. When expert teams are not located near the infrastructure, they may need to travel considerable distances to access the site and perform their tasks. Inspections frequently require close examination of critical components, which can involve physically demanding and high-risk activities, such as climbing walls or pillars or descending from heights using abseiling techniques. The proposed method provides a practical solution that allows engineers to conduct regular infrastructure inspections remotely, eliminating the need for site visits and reducing exposure to hazardous tasks. The system benefits not only inspection companies but also public and private entities responsible for infrastructure management and visualization as well as the movie industry. Besides that, access to relevant infrastructure can be done faster especially in cases of emergency if needed.

7. REFERENCES

[1] The Economist. 2025. Los Angeles – Mixed reality technology. Retrieved Jan 20, 2025 from https://innovation-matters.economist.com/losangeles/mixed-reality-technology

[2] Ismael Colomina and Pete Molina. 2014. Unmanned aerial systems for photogrammetry and remote sensing: A review. ISPRS J. Photogrammetry and Remote Sensing 92, (June 2014), 79–97. http://dx.doi.org/10.1016/j.isprsjprs.2014.02.013

[3] N.M. Noor, A. Abdullah, and M. Hashim. 2018. Remote sensing UAV/drones and its applications for urban areas: a review. IOP Conf. Series: Earth and Environmental Science 169. http://dx.doi.org/10.1088/1755-1315/169/1/012003

[4] F. Nex, C. Armenakis, M. Cramer, D.A. Cucci, M. Gerke, E. Honkavaara, A. Kukko, C. Persello, and J. Skaloud. 2022. UAV in the advent of twenties: Where we stand and what is next. Review Article. ISPRS J of Photogrammetry and Remote Sensing, 184, 215-242. https://doi.org/10.1016/j.isprsjprs.2021.12.006

[5] S.I. Granshaw. 2018. RPV, UAV, UAS, RPAS … OR JUST DRONE? Editorial. The Photogr. Record, 33 (162), 160-170. https://doi.org/10.1111/phor.12244

[6] G. Forlani, E. Dall'Asta, F. Diotri, U.M. Cella, R. Roncella and M. Santise. 2018. Quality assessment of DSMs produced from UAV flights georeferenced with on-board RTK positioning. Remote Sensing, 10, 311. https://doi.org/10.3390/rs10020311

[7] D. Cao, B. Zhang, X. Zhang, L. Yin, X. Man 2023. Optimization methods on dynamic monitoring of mineral reserves for open pit mine based on UAV oblique photogrammetry. Measurement, 207, 112364. https://doi.org/10.1016/j.measurement.2022.112364

[8] J. Bowman, L. Yang, O. Thomas, J. Kirk, A. Duncan, D. Hughes and S. Meade. 2023. UAS Edge Computing of Energy Infrastructure Damage Assessment. Photogr. Eng. and Remote Sensing, 89, 2, 79–87. https://doi.org/10.14358/PERS.22-00087R2

[9] C. Li, F. Zhu, B. Guo, Z. Wang, X. Jiang, J. Wang and X. Liao. 2022. Power Line Extraction and Obstacle Inspection of Unmanned Aerial Vehicle Oblique Images Constrained by the Vertical Plane. The Photogr. Record, 37, 306-332. https://doi.org/10.1111/phor.12422

[10] S. Meng, Z. Gao, Y. Zhou, B. He and A. Djerrad. 2022. Real-time automatic crack detection method based on drone. Compt. Aided Civil & Infrastructure Eng., 1–24. https://doi.org/10.1111/mice.12918

[11] X. Zhong, X. Peng, S. Yan, M. Shen, Y. Zhai. 2018. Assessment of the feasibility of detecting concrete cracks in images acquired by unmanned aerial vehicles. Automation in Const., 89, 49-57. https://doi.org/10.1016/j.autcon.2018.01.005

[12] Y.F. Liu, X. Nie, J.S. Fan and X.G. Liu. 2020. Image-based crack assessment of bridge piers using unmanned aerial vehicles and three-dimensional scene reconstruction. Comput. Aided Civ. Inf., 35, 511–529. https://doi.org/10.1111/mice.12501

[13] T. Yamane, P.-J. Chun, J. Dang, and R. Honda. 2023. Recording of bridge damage areas by 3D integration of multiple images and reduction of the variability in detected results. Comput. Aided Civ. & Infr. Eng., 1–17. https://doi.org/10.1111/mice.12971

[14] D. Germanese, G.R. Leone, D. Moroni, M.A. Pascali, M. Tampucci. 2018. Long-Term Monitoring of Crack Patterns in Historic Structures Using UAVs and Planar Markers: A Preliminary Study. J. Imaging, 4, 99. https://doi.org/10.3390/jimaging4080099

[15] Lei Xiao, Salah Nouri, Joel Hegland, Alberto Garcia Garcia, and Douglas Lanman. 2022. NeuralPassthrough: Learned Real-Time View Synthesis for VR. In ACM SIGGRAPH 2022 Conference Proceedings (SIGGRAPH '22). Association for Computing Machinery, New York, NY, USA, Article 40, 1–9. https://doi.org/10.1145/3528233.3530701

[16] T. Marrinan and M. E. Papka. 2021. Real-Time Omnidirectional Stereo Rendering: Generating 360° Surround-View Panoramic Images for Comfortable Immersive Viewing. IEEE Transactions on Visualization and Computer Graphics, 27 (May 2021), 5, 2587-2596. DOI: 10.1109/TVCG.2021.3067780.