See What We See - Sharing Mixed Reality Experiences with WebRTC

Richard Barnett
Arvizio, Inc.

Jonathan Reeves
Arvizio, Inc.

Borys Vorobyov
Arvizio, Inc.

ABSTRACT

Mixed Reality (MR) combines concepts from both AR and VR placing virtual objects in positional space around you. Using spatial anchoring techniques models can be 'placed' on a conference table, or in free space, and then observers can walk around the model. Sharing this experience with others who are not wearing MR headsets requires new tools for capturing and recording the experience for viewing on local monitors or using WebRTC to extend the mixed reality feed over the Internet to remote participants.

Keywords: Mixed Reality, Spatial Mapping, Real-Time Conferencing, 3D.

Index Terms: Mixed / Augmented Reality, Collaboration, Computer Conferencing, Video Conferencing, Rendering, Image Compression, Shape Modeling and Visualization.

1 INTRODUCTION

MR Studio is an advanced software platform developed by Arvizio, Inc. for augmented and mixed reality experiences offering real time collaboration and enhanced visualization utilizing the Microsoft HoloLens and other leading mixed reality devices. The platform is designed to serve, and bring new levels of efficiency to a variety of industry verticals including AEC (Architecture, Engineering and Construction), healthcare, industrial IoT (Internet of Things), enterprise training and general education. The goal of the development is to offer real-time mixed reality visualization and collaboration, both in a local setting and across locations by leveraging WebRTC as an underlying communications technology.

Mixed Reality (MR) combines concepts from both AR and VR but places 2D & 3D virtual objects in positional space around you. For example, using mixed reality you are able to walk up to virtual objects and interact. Using spatial anchoring techniques models can be 'placed' on a conference table, or in free space, and then observers can walk around the model viewing it from all sides without sacrificing awareness of your physical surroundings.

Augmented Reality (AR) is a widely used component of a mixed reality experience. AR techniques can be used to recognize identification markers and trigger the display of data specific to the object for example, or to display information from the real time data feeds associated with a specific piece of machinery. Mixed reality devices use SLAM (Simultaneous Localization and Mapping) techniques to track ones position which then can be used as a trigger for AR displays. In fact, mixed reality can be understood as the seamless integration of augmented reality with your perception of the real world.

Richard Barnett Email: rbarnett@arvizio.io
Jonathan Reeves Email: jreeves@arvizio.io
Borys Vorobyov Email: bvorobyov@arvizio.io

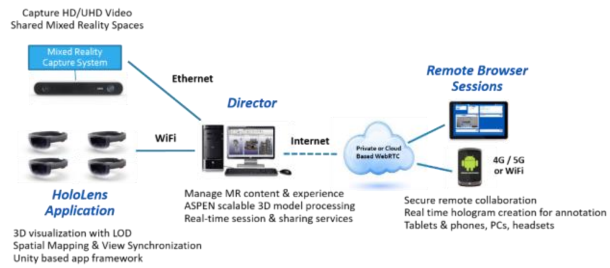

Figure 1. Arvizio MR Studio Architecture.

2 3D MODEL PROCESSING

Applications increasingly require the use of large 3D model datasets that are beyond the processing capabilities of AR/MR headset devices such as smart glasses, smartphones and tablets. While the processing power of 'Stand-Alone' devices has increased substantially, there is still a significant gap between the CPU and GPU processing power available in such devices and the power of desktop workstations. Optimizing the data hierarchy and controlling the interaction between the remote devices, local edge computing servers and the cloud, requires a Well-Designed data management architecture that minimizes latency and shares the processing burden.

2.1 Level of Detail Processing

Level of Detail (LOD) processing is a familiar concept but mixed reality devices change the requirements. For example, the level of detail may need to change dynamically as the user approaches an object - the closer to the object, the greater the resolution required. This requires rapid indexing of data and the ability to traverse the LOD database in real time. In addition, there are a variety of 3D data formats including 3D textured meshes, spatial data such as LiDAR (Light Detection And Ranging) point clouds, 3D volumetric models and raster data representations. Each has their own unique characteristics in terms of LOD processing techniques, file formats and approaches used to render the data on the mixed reality headset devices.

2.2 Spatial Mapping

A unique spatial data concept associated with some mixed reality devices is spatial mapping. HoloLens, for example, has multiple sensors that allow the device's software to construct a spatial model of the room in which the wearer is situated. These Device-Generated spatial maps are used to aid in the tracking of head movement and offer the ability to 'anchor' a 3D model to a particular location in the room. In the future, spatial maps will become valuable data in their own right. Spatial maps of many rooms can be stored in a database then, on entering a particular room, the spatial map can be reloaded along with virtual display panels showing information pertinent to the particular room.

One of the key objectives has been to operate as a data server thereby providing data dynamically with the correct level of detail (LOD) based on the needs of the viewing experience. The platform then serves to augment the local processing of the attached headsets and can serve optimized versions of the data as required. This allows very large models in the realm of many gigabytes of data, millions of polygons, or billions of points, to be displayed on the headset by serving the smaller segments of the model as required.

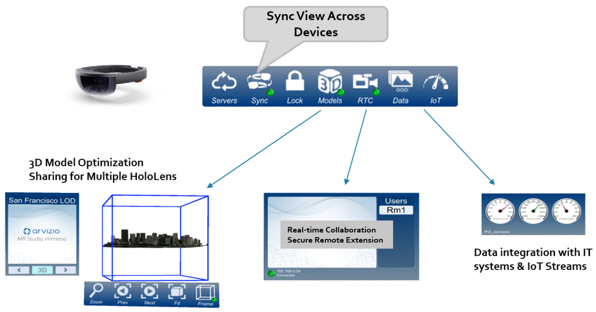

Figure 2. MR Glasses Application View.

The 3D rendering and visualization technology used for engineering use cases, such as processing CAD models, can also be applied to a number of areas in the medical field. In radiology for example, the primary data formats are known as DICOM. CT and MRI scans process and store data in the form of sequential 2D slices. A scan may consist of several hundred slices which are usually viewed on 2D monitors. Development of 3D large model processing technology is essential to construct a 3D model from the many individual slices and automatically push the model to the HoloLens device for visualization. This provides an entirely different viewing experience and opens up new avenues of training, improved rehearsal of complex surgical procedures and improved interaction with the patient. In other industries, such as engineering and construction, the ability to view and interact with a virtual 3D model created from CAD drawings is important for many reasons. It serves as an essential, and cost saving, method to present the finished product (without the need for a physical model), identify potential flaws early in the process, get all involved parties on the same page and, provide a positive end-customer interaction. Part of the collaborative experience includes the ability to direct a participant's focus to a specific object or feature of the 3D model being viewed. This requires that everyone in the experience have the ability to highlight and/or annotate the feature of interest.

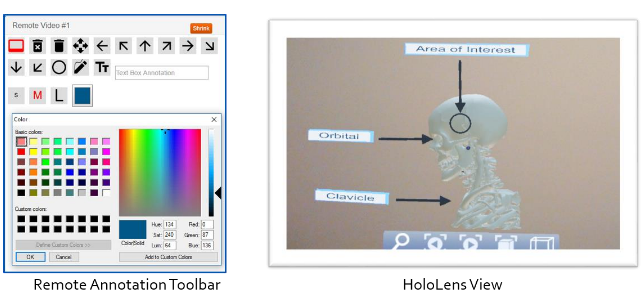

Figure 4. Remote MR Annotation in a Web Browser.

The nature of mixed reality tends to expose us to adjacent fields across industries. The common thread however, is the need to visualize complex 3D information.

3 Holographic Collaboration

Mixed reality devices often use game engine platforms for the development of the applications that operate on the device. Unity 3D is the most commonly used; it is a powerful platform but it does have limitations when dealing with spatial data and real-time network communications. The enhancement of the native capabilities of Unity 3D with native plugins optimized for spatial data processing, allows cross platform development with improved performance.

Holographic collaboration allows 3D images to be viewed by many team members, each with their own unique viewing perspective. The seamless integration of holographic computing and video communications creates an exceptional experience for sharing real world visual information and augmented reality data among team members.

For a solution to truly offer a collaborative experience it must include software that runs on the MR headset as well as software that runs on a connected laptop, desktop computer or server to feed content to multiple devices while allowing the operator to control the experience. An additional aspect of a rich mixed reality experience is sharing live video streams. The WebRTC protocols and application interfaces are ideal for enabling bandwidth efficient video calls in a private, secure and fully integrated environment. Through video and audio sharing, subject matter experts can observe live video and mixed reality content remotely and participate in the interactive experience.

In many instances, all of the interested parties may not be in the same location creating a need for a real-time collaboration capability to share live video and audio feeds from the mixed reality device with remote participants, such as remote specialists. A truly interactive and collaborative mixed reality solution will allow the remote specialist to participate using a two-way live video feed and view the overlaid holographic models using a laptop, tablet or even a smartphone device.

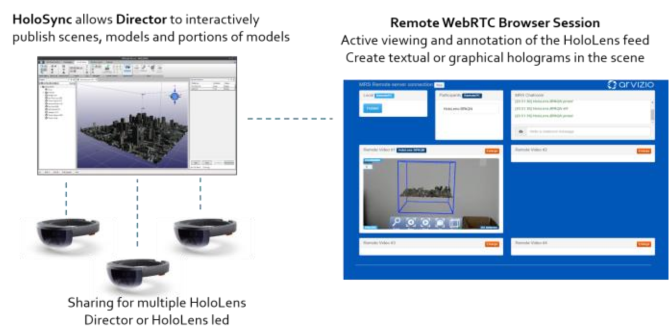

Figure 3. HoloSync Coordinated Experience.

Holographic collaboration allows 3D images to be viewed by many team members, each with their own unique viewing perspective. The seamless integration of holographic computing and video communications creates an exceptional experience for sharing real world visual information and augmented reality data among team members.

4 HoloLens Spectator View

Spectator View is a Microsoft developed tool set allowing others to "see" what the HoloLens wearer is viewing without the need to put on the headset. Spectator View combines an open source software solution, Unity Editor, a dedicated HoloLens, high quality digital SLR and video capture hardware to provide a means of recording mixed reality scenes and sharing the scene with an audience.

While Spectator View is an important development in the ability to demonstrate the mixed reality experience, the cost to build the rig and complexity of the set up can be prohibitive for many situations. The cost of a Spectator View configuration is around $6000 (USD) for the additional HoloLens and DSLR camera alone which is too expensive for widespread usage. The following figures illustrate the use and configuration of Spectator View.

Figure 5. Spectator View Rig. Source: Microsoft

Figure 6. Spectator View Hardware. Source: Microsoft

5 MR STUDIO VIEWPOINT

The authors are evaluating an alternative that replaces the need for a dedicated HoloLens and DSLR with low cost 3D depth sensing cameras. This solution is approximately $1/10^{th}$ of the cost of the Spectator View configuration and does not require a dedicated HoloLens. When coupled with WebRTC-based delivery, it allows a conference session to be efficiently shared with local and remote audiences alike. This is an important breakthrough in the ability to share the entire mixed reality experience with an audience. The configuration associated with MR Studio ViewPoint is illustrated in Figure 7.

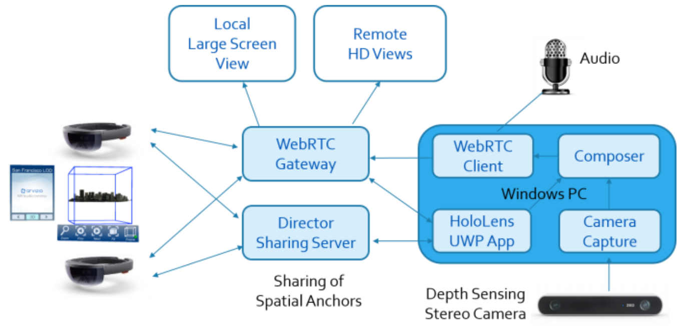

Figure 7. MR Studio ViewPoint.

Figure 8. Prototype ViewPoint Calibration Rig.

A depth-sensing camera (for example, the Stereolabs ZED) is used to generate an overview of the Real-World space. The application that runs on the MR device also runs on the ViewPoint PC. The Unity camera textures and other relevant data is passed from the application into the Composer, where video from the depth sensing camera is combined to form a mixed reality stream. Finally, the MR stream is passed to the WebRTC client for distribution that makes it possible to consume the mixed reality stream at any location using devices that can run a web browser with WebRTC support (desktops, laptops, mobile devices, etc.).

An essential requirement of such a system is that the Unity based application on the ViewPoint PC is set to correctly represent the position and orientation of objects in the real space and correctly matches the scenes seen by the HoloLens users. Spectator View solves this by attaching a HoloLens permanently to the camera. ViewPoint however, uses an initial calibration mode where a HoloLens is attached to the depth sensing camera but, once the calibration process is complete, the HoloLens can be detached and used.

The calibration process uses feature matching between the field of view as seen by the ViewPoint camera and that seen by the attached HoloLens camera. The calibration rig ensures that they are viewing approximately the same scene to assist with feature matching but the alignment does not need to be exact. Feature matching allows the system to determine two important pieces of information: the actual field of view of the ViewPoint camera and the pose offset between it and the HoloLens camera. Knowing the pose of the HoloLens in the space allows the ViewPoint software to compute the information needed to correctly position the Unity coordinate system.

Once calibration has been completed, the HoloLens and ViewPoint camera can move independently. Depth sensing cameras such as the ZED are able to track their pose in real space. This means that the ViewPoint camera can perform pans, tilts and dollies while maintaining a correct mixed reality stream.

In the future, ViewPoint may also make use of non-depth sensing cameras and other mixed reality devices. However, in this case, the camera must either remain fixed in place after calibration or be augmented with other sources of pose information such as an attached IMU or LIDAR scanning device.

6 Conclusion

The ability to share and record Mixed Reality (MR) experiences, both locally and at remote locations, is an essential ingredient for commercial use cases. Sharing this experience with others who are not wearing MR headsets requires new tools, and Arvizio has successfully demonstrated the use of WebRTC as an underlying fabric to extend the mixed reality data feed over the Internet. MR Studio is an advanced software platform developed by Arvizio, Inc. for augmented and mixed reality experiences offering real time collaboration and enhanced visualization utilizing the Microsoft HoloLens and also other emerging mixed reality devices. Our platform brings new levels of efficiency to a variety of industry verticals, and the development of MR Studio ViewPoint will serve to further enhance the quality of the experience.

References

[1] https://developer.microsoft.com/en-us/windows/mixed-reality/spectator_view

[2] https://developer.microsoft.com/en-us/windows/mixed-reality/mixed_reality_capture_for_developers

Richard Barnett, Vice President, Solutions Architecture, is a recognized leader in the design and implementation of advanced technologies including networking, IoT and robotics platforms. Richard co-founded and served as CTO of Mangrove Systems, served as VP Systems Architecture at Sirocco Systems (acquired by Sycamore). Previously, Richard founded Netcomm (acquired by GDC). Richard earned an M.S. in Telecommunications from the University of Essex and served as an Operating Partner and Chief Technologist at Bessemer Venture Partners. Richard has authored textbooks and several papers relating to the design and application of technology.

Jonathan Reeves, CEO, is a recognized leader and serial entrepreneur experienced in the development of forward looking technology business ventures. Prior to co-founding Arvizio, Jonathan served as Chairman at CloudLink Technologies, (acquired by EMC), founded Mangrove Systems, (acquired by Carrier Access), Sirocco Systems (acquired by Sycamore Networks) and Sahara Networks, (acquired by Cascade communications). Jonathan has served as an Operating Partner with Bessemer Venture Partners and previously served on the Board of Trustees at Quinnipiac University.

Borys Vorobyov, CTO, has many years of hands-on experience in software development in spatial data analysis, LiDAR spatial data processing and plant control systems. Prior to co-founding Arvizio, Borys was the founder and CEO of SightPower, a company focused on large scale spatial data analysis and 3D visualization for the mining industry and BIM. Borys co-founded Ambercore Software and served as CTO leading the development of LiDAR point cloud analytics solutions used in a range of GIS applications. As Lead Scientist for Kharkov National University, Dr. Vorobyov led numerous projects, including informational support and modeling for oil and gas exploration and plant control automation systems.