

Case Study

High-Speed Cameras Break Down Barriers for Real-Time Mixed Reality

Imagine being able to stand next to your favorite performer during a once-in-a-lifetime concert — or having reserved seating at the 50-yard line for every game or the chance to “run” beside star players as they charge the goal during a league championship. Now, imagine being able to have these experiences while sitting in your living room, commuting home from a busy day at work, or playing the latest multiplayer game.

You-are-there experiences are what immersive media promise to deliver with real-time mixed reality (MR). This new format uses volumetric video data to create three-dimensional (3D) images as an event is occurring (Figure 1). Further, multiple people can view the images from different angles on a range of devices.



Capturing Reality in 3D Is Hard

Media companies have been early adopters of technology formats such as 360-degree video, virtual reality (VR), augmented reality (AR), and MR.

Regular, as opposed to real time, MR blends physical and digital objects in 3D and is usually produced in dedicated spaces with green screens and hundreds of precisely calibrated cameras. Processing the massive amounts of volumetric data captured in each scene requires hours or even days of postproduction time.

Trying to do MR in real time has proven even more technically and economically challenging for content developers and, so far, has made the format impractical.

“Capturing and synchronizing high-resolution, high-frame-rate video from a large number of cameras by itself is simple enough for us,” said John Ilett, CEO and founder of Emergent Vision Technologies, the premier manufacturer of high-speed imaging cameras. “Processing this video in real-time in live venues does indeed have its challenges.”



Figure 1: Condense Reality uses high-speed GigE Vision cameras from Emergent Vision Technologies to produce immersive mixed reality experiences.



Deep Learning Needs an Assist

One startup thought it had a strategy for overcoming those issues. Condense Reality, a volumetric video company, had a plan for capturing images, reconstructing scenes, and streaming MR content at multiple resolutions to end-user devices. From start to finish, each frame in a live stream would require only milliseconds to complete.

“Our software calculates the size and shape of objects in the scene,” said Condense Reality CEO Nick Fellingham (Figure 2). “If there are any objects the cameras cannot see, the software uses deep learning to fill in the blanks and throw out what isn’t needed, and then streams 3D motion images to phones, tablets, computers, game consoles, headsets, and smart TVs and glasses.”

But there was a hitch. For the software to work in real-world applications, Fellingham needed a high-resolution, high-frame-rate camera that content creators could set up easily in a sports stadium, concert venue, or remote location. The company tested cameras, but the models used severely limited data throughput and the cable distance between the cameras and the system’s data processing unit. To move forward, Condense Reality needed a broadcast-quality camera that could handle volumetric data at high speeds.



Figure 2: Condense Reality CEO Nick Fellingham stands inside one of the company’s volumetric capture setups.



High-Speed Cameras Deliver

In 2020, Fellingham learned that Emergent Vision Technologies was releasing several new cameras with high-resolution image sensors. These cameras included models with [SFP28-25GigE](#), [QSFP28-50GigE](#), and [QSFP28-100GigE](#) interface options, all of which offer cabling options to cover any length.

“Our cameras deliver quality images at high speeds and high data rates,” said Ilett. “They capitalize on advances in sensor technology and incorporate firmware we developed so the cameras can achieve the sensor’s full frame rate.”

The images in an MR experience should be captured at an extremely high frame rate and resolution. With the new cameras, Fellingham was able to assemble a commercially viable system. “High-speed GigE cameras are what we need to get the data off the cameras quickly and stream it,” he noted.

High-speed capture is particularly important for sports, where exciting action often occurs in the literal blink-of-an-eye. When capturing a golf swing, for example, a camera with a frame rate of 30 fps is likely to only “see” the beginning and end of the swing, which significantly reduces the quality of the volumetric content.

“We are not using these cameras for inspecting parts in a factory; we are using them to create entertainment experiences,” said Fellingham. “When the speed [fps] increases, the quality increases for fast-moving action, the output we generate is better, and the experience improves overall.”

Bigger Capture Areas Are on the Horizon

Condense Reality serves as a system integrator for customer projects. A standard system uses 32 cameras, one high-speed network switch from Mellanox, and a single graphics processing unit (GPU) from NVIDIA to cover a 7-meter by 7-meter capture. The company worked with Emergent Vision Technologies to put together the optimal system for volumetric capture.

“We don’t necessarily want to be committed to very specific hardware configurations, but by working with the Emergent team and testing different components, we’ve found that NVIDIA and Mellanox work best for us,” said Fellingham.



Along with implementing its technology, the company is working to increase the capture area for MR while maintaining throughput and quality.

“When you start to get bigger than a 10 by 10 meter area, 4K cameras don’t cut it,” Fellingham said. “When our algorithms improve, we will go bigger.”

The new Emergent Vision Technologies cameras are integral to this work. With models supporting [up to 600 fps](#) at 5120 x 4096 resolution and interface options ranging up to 100 GigE, Fellingham has not had to worry about caps on camera resolution, data rates, or frame rates. Those advantages mean that Condense Reality is well positioned to deliver even better content and user experiences.

Software: The Secret Sauce

Condense Reality’s software is a completely proprietary offering that creates a 3D mesh comprised of hundreds of thousands of polygons. The nodes placed on an object represent the surface of the object being captured. Then the software “paints” the mesh with data acquired by the cameras. This data then uses deep learning to estimate the parts of the object that the cameras did not cover. Compression algorithms then reduce the mesh down to as small a size as possible for every frame, said Fellingham.

“The software takes all of this data and turns it into a 3D model at an extremely quick pace, and it can only do this because of well-optimized algorithms, neural networks, and the NVIDIA GPUs,” he said. “While most neural networks are based in TensorFlow, some of the ones we use in the system need to run very fast, so they’re written directly for the GPU.”

He added, “Our neural networks solve very specific problems, which helps when optimizing them for speed. We don’t deploy a huge black box that performs a ton of inference, as this would be very hard to optimize.”



To finish the process, data is sent to Condense Reality’s cloud-based distribution platform, which takes the data and puts it into a variable bit rate so that the stream is different depending on the user device. Playback happens inside a game engine, which allows customers to build custom experiences, which could be VR or AR, around the volumetric video. Because Condense Reality systems support game engines, their content can also be streamed into existing game worlds owned by other companies. Currently the software supports Unity and Unreal game engines, but the company plans to build plug-ins for its software for any new game engines that emerge in the future.

“These engines can barely be called game engines anymore, as they’re really interactive, 3D tools,” said Fellingham. “We route the real-world content into these tools to provide customers with photorealistic 3D, interactive experiences.”

FOR FURTHER INFORMATION

Emergent Vision Technologies’ High-Speed Cameras: www.emergentvisiontec.com/products/area-scan-cameras

Condense Reality: www.condensereality.com

Any Questions?

866-780-6082

sales@emergentvisiontec.com

<https://emergentvisiontec.com>