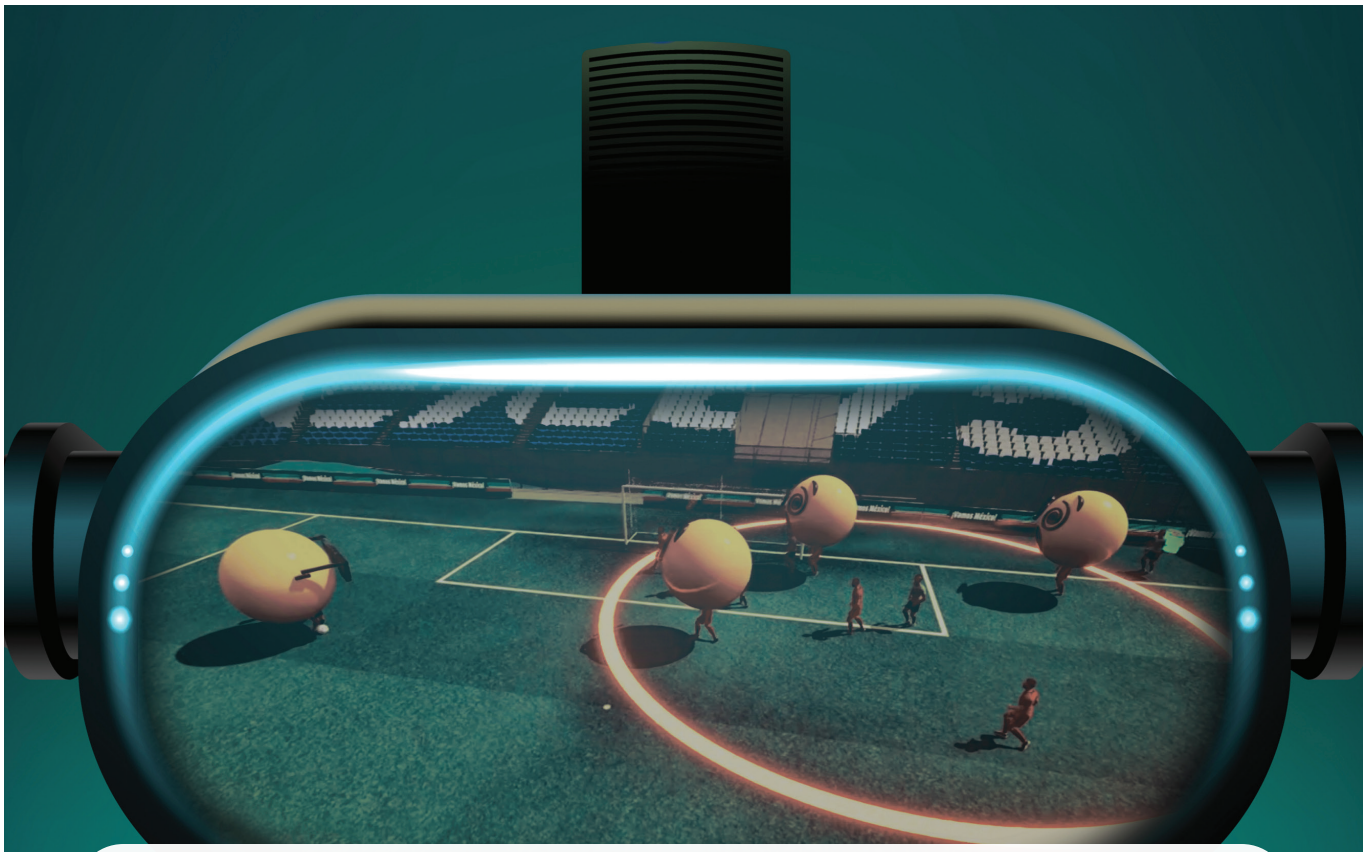


Introduction to Immersive and Volumetric Media

V. MICHAEL BOVE, JR.

The engineering and design behind immersive media experiences aim to minimize or eliminate the perceived separation between the audience and the content, providing the sense that the experience is happening in the viewer's space rather than within a frame or on the other side of a window. Multiple aspects of content creation and reproduction can contribute to a viewer's feeling of immersion and involvement.

The spatial and angular extent of the display are powerful visual cues for supporting the impression of being in the content space, and extending the light field of the imagery beyond the visual angle of the viewer helps close the gap between the viewer's physical environment and the scene environment. Projection systems have aimed toward this goal for many years, from the first Zeiss planetariums in the 1920s through such processes as

1545-0279/24©2024SMPTE

CinemaScope, VistaVision, Cinerama, the Walt Disney Company's Circle-Vision 360°, and IMAX/Omnimax. The current progression of ever-larger flat-panel television screens is a domestic pursuit of the same effect. Wearable display devices (which trace their lineage back to the stereoscopes of the early 1800s, predating even still photography) provide another mechanism for creating immersive light spaces, and the effectiveness of their user experience is closely tied to the angle of view they can fill with imagery.

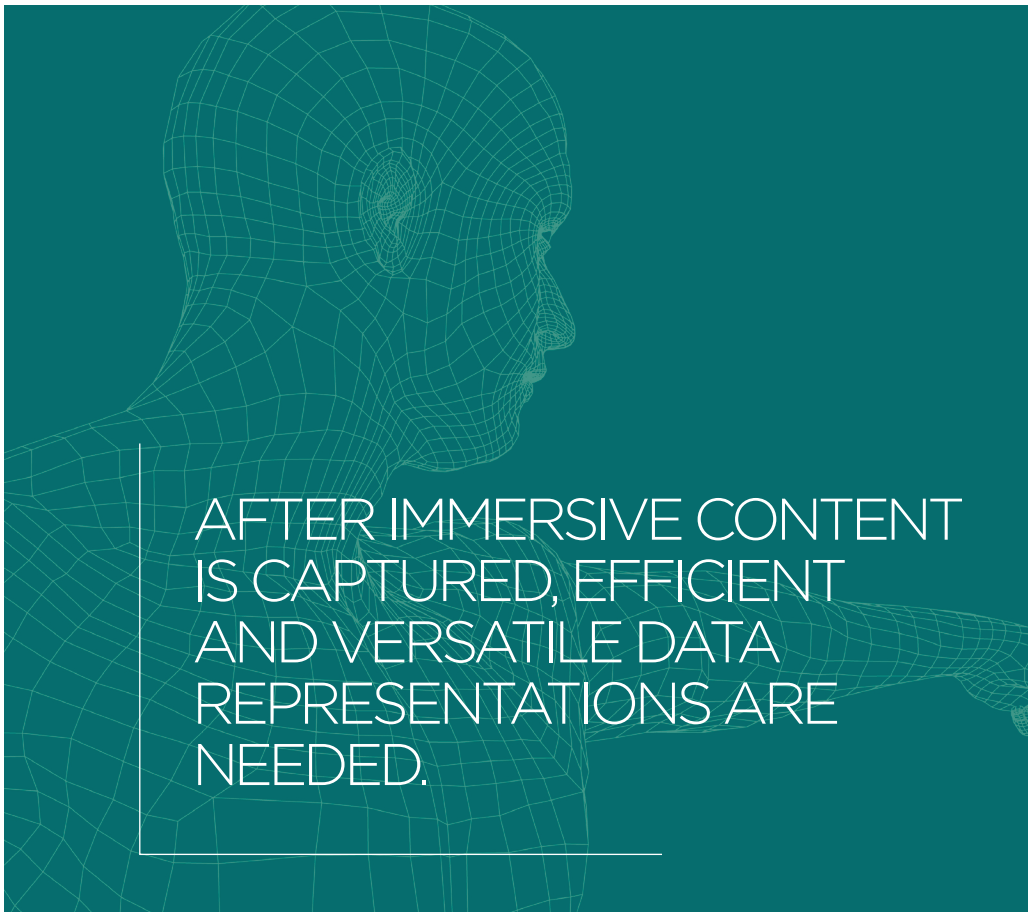
Reproducing such experiences not only pushes the limits of display technology but also demands the capture and transport of far more information than traditional television or cinema, necessitating new engineering developments throughout the entire pipeline and new creative approaches. Unlike displays of the past, immersive displays now are tightly integrated with previously unimaginable amounts of computation, opening equally unimaginable technical and expressive potential, while representations of media are now becoming efficient descriptions of dynamic 3D scene objects and properties rather than simply recordings of optical signals.

Navigability, a capability of some immersive media content that has become increasingly practical in recent years, means that the viewer position and viewing direction can be varied naturally, allowing exploration of the content space. The most flexible and realistic navigable experiences are enabled by volumetric capture, in which source material is recorded or authored in three dimensions rather than as 2D images and, through computer graphics techniques, is rendered on the fly for viewing from various virtual viewpoints. Several generations of consumers have grown up with navigability through video games, and extending that degree of freedom to other genres, such as real-world sporting events through game engines, is now an intuitive concept for audiences. Volumetric capture, as a bonus, provides the capability to drive 3D displays, which can further enhance immersiveness and encourage exploration of the space; it also creates highly versatile representations of scene elements that support effects like 3D compositing and relighting.

In summary, providing an immersive visual experience entails the capture or creation of high-visual-fidelity source material, the efficient distribution of that content with acceptable quality of service (QoS), and finally the reproduction of the content for the audience in an immersive way. SMPTE members have, from the beginnings of the Society to the present day, worked at the forefront of expanding the audience experience, including in the realm of immersive media; exactly five years ago, the *Journal* issue was also focused on this topic. The technological landscape has changed remarkably since then through such developments as a dramatic increase in the processing available in consumer devices–particularly mobile ones, the deployment of much faster distribution infrastructure (e.g., 5G), and advances in algorithms and tools for capturing and working with content. This issue features papers exploring some of the latest advances in each of those areas.

The paper "Advanced Volumetric Video Format for Enhancing Photo-Realistic Lighting Reproduction," by Yuma Wakahara, *et al*., presents a novel volumetric capture system coupled to a powerful rendering technique, with particular emphasis on enabling re-illumination of the scene elements.

In "Using Metahumans in Large-Scale Sports Production," Doris Retfalvi also describes a system extend-



ing to both ends of the chain, in a case study of how a sporting event was recreated in 3D by means of digital doubles of athletes and a game engine.

5G network deployment for extended reality (XR) experiences involves more than just data transport, and in "The Evolution of Delivering Immersive Media over 5G Cloud," Mauricio Aracena, *et al.* explores how edge computing is integrated into the 5G ecosystem to support powerful, responsive immersive experiences without being limited by the hardware capabilities of user devices.

After immersive content is captured, efficient and versatile data representations are needed. Thomas Stockhammar reviews the current state of activities by the MPEG Systems Working Group on Scene Description in "MPEG-I Scene Description: A Dynamic Scene Description Framework for Immersive Media."

Navigable media imposes critical bandwidth and timing requirements on delivery networks, and user experience is adversely impacted by packet loss, jitter, and delay. "Optimizing the Virtual Reality Experience Using a 360° View of Client and Network Data," by S. Mishra, *et al.* discusses experiments by the Streaming Video Technology Alliance's Immersive Video Working Group to address the engineering challenges of delivering immersive content to the home.

Another topic currently receiving great attention and undergoing rapid development is artificial intelligence (AI) tools for media creation. While not specifically examining immersive experiences, "Efficient Media Production and Management with AI Assistants: A Multi-Domain Exploration," by Rob Gonsalves, *et al.* provides a look into how machine learning (ML) systems can collaborate with human content makers and amplify their creativity and productivity, rather than replacing them.

While this issue is specially focused on the topic, the pursuit of immersiveness continues to be a driver of much of the most cutting-edge engineering and creative innovation in digital media, and the *Journal* encourages readers to submit papers advancing the science and art in this area at any time.

About the Author

V. Michael Bove, Jr. has published over 100 articles and conference papers in video, optics, user interface, product design, 3D capture, and wireless networks. He holds over 30 patents.

DOI: 10.5594/JMI.2024.ZMQJ6921
Date of publication: 15 May 2024

1545-0279/24©2024SMPTE