DEFINING THE XR EXPERIENCE: ENABLING THE IMMERSIVITY ECOSYSTEM

ABiresearch. THE TECH INTELLIGENCE EXPERTS™ interdigital.

Eric Abbruzzese: Research Director

CONTENTS

WHAT IS XR AND IMMERSIVE CONTENT?

Extended Reality (XR) is a broad term encapsulating a spectrum of Augmented Reality (AR), Virtual Reality (VR), and Mixed Reality (MR) experiences that capture audio, video, haptic, and kinesthetic data into one scene representation to seamlessly blend interactions in the virtual and physical world. Whether it be AR leveraging transparent displays for seamless data visualization, VR entirely replacing physical with virtual, or MR somewhere in between, XR has increasingly seen adoption across markets.

With a novel, yet challenging form factor in Head-Mounted Displays (HMDs), XR can be difficult to implement depending on the use case and target user. Today, however, new efficiencies in encoding techniques, alongside new media types and functionalities that support high-quality and realistic media streaming and interactive services, are converging with enhanced network and device processing capabilities to enable an array of groundbreaking XR-enabled immersive experiences. This enables them to sidestep some of the local hardware challenges. XR-enabling capabilities hold significant potential to transform consumer experiences, services, and entire industry verticals, and standardization is an important step toward success.

This paper outlines the foundational technologies and metrics that define XR experiences, introduces the current standards study items and research resources dedicated to enabling XR, and highlights some of the challenges that will define the next stage of innovation and standardization for XR-enabled immersive experiences.

DEVICES AND EXPERIENCES

The application of combining physical and digital worlds is not new, but the state of XR hardware, software, and services enabling these applications has reached a turning point for capability versus drawbacks. Price has always been a barrier for new technology adoption, and while there is still room for price decreases, hardware Average Selling Prices (ASPs) have been trending downward, with increasing competition under US$1,000 and even under US$500. Today, the number of active developers and content houses targeting XR services and applications is steadily increasing.

XR Device Examples

AR
RealWear Navigator 500

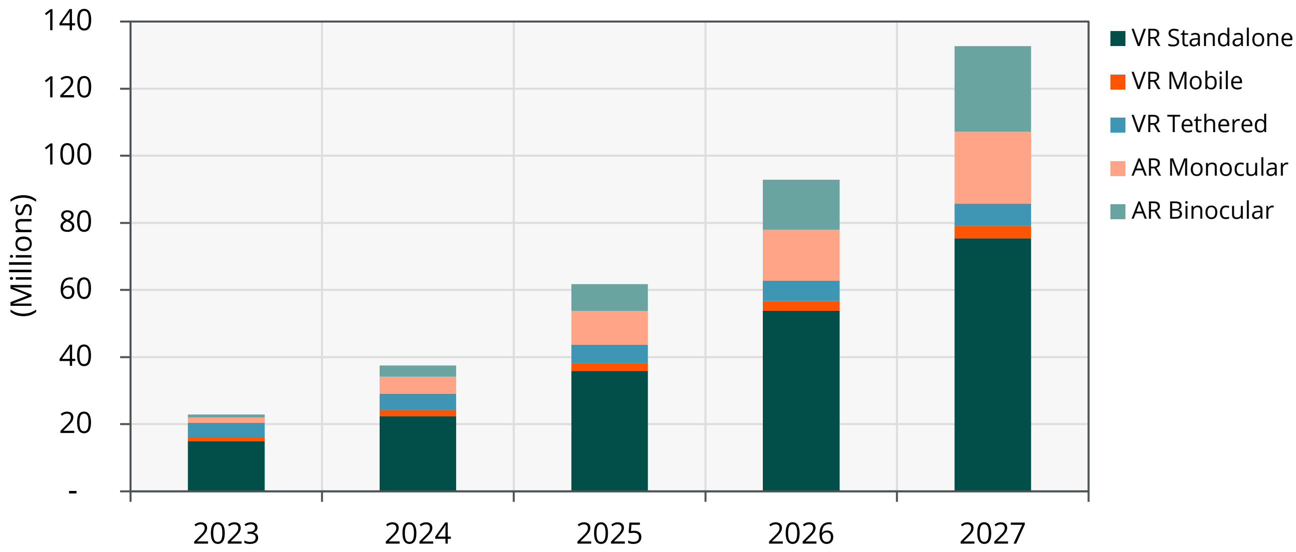
MR
Microsoft HoloLens 2

VR
Apple Vision Pro

Some segmentation still exists within these device categories. VR, for instance, can be broken out broadly into standalone, tethered, and mobile-based solutions. Standalone devices contain all hardware required to process and display VR content on the device, requiring no wired or wireless connection to another device. Tethered VR uses a wireless or wired tether to a host device that often does computation, enabling greater performance, while maintaining an acceptable form factor. Mobile-based VR once was the largest segment, with devices like Google Daydream and Samsung Gear, but this segment is now small relative to the standalone and tethered segments.

Although mobile-based VR represents a relatively small segment of the overall market, the benefits that are offered by mobile-based VR are too large to ignore. The ability to connect devices anywhere an Internet connection is available opens the door to an immeasurable number of new use cases. Mobile-based VR platforms are generally well understood, include highly supported hardware and connectivity options, and have well-established development communities. Even so, standalone VR will likely continue to be the dominant VR device type over the next few years.

XR Shipments by Form Factor



The success of the standalone segment is due mostly to a combination of affordability (Meta Quest 3 will cost US$500, and the last generation Meta Quest 2 will cost US$300) and ease of use. There are some affordability exceptions with high-end devices, but the ease of use and intuitive user experience remains. Affordability is not a universal truth for standalone VR, with high-end devices like Apple's Vision Pro aiming to provide a best-in-class experience. There is a balancing act with XR compute; ease of use comes with costs in battery life, user comfort, and outright performance capabilities. Streaming content from the edge can alleviate some or all those costs, enabling lower cost and more ergonomic HMDs, but it presents its own obstacles in high data rate and low latency requirements.

Today, the enterprise market is accelerating VR adoption significantly alongside AR, while the consumer market is waiting for competitive smart glasses at affordable price points. Smart glasses have proven difficult to develop an acceptable package of hardware at an acceptable cost; transparent display quality, battery life, and design/form factor have proven significant challenges. VR headsets that pass video live through front-facing cameras are filling this gap for now, with a similar result for true transparent displays, albeit with a sacrifice to pass-through quality and user awareness.

WHY DO WE NEED XR?

To be deemed a necessity, XR must show objective and irrefutable value that outweighs any potential downsides, while also surpassing a minimum level of user experience to support the targeted workflow—user comfort, battery life, and device capability must all align with the use case needs. In the enterprise sector, this has already been accomplished with numerous case studies and Proofs of Concept (PoC) now scaling up to more users, devices, and use cases. Enterprises were the first to adopt AR smart glasses, while consumers were first to adopt VR. This was dictated mostly by the valuable use cases in verticals; enterprises saw immense value from hands-free device usage for remote assistance, work instruction, and training. VR presented an entirely unique experience for video and gaming content, and continues to see steady growth in that space.

Worker enablement is the key for enterprises and XR. Spurred by COVID-19 lockdowns and movement toward hybrid workforces, companies have recognized that the benefits of well-supported hybrid workforces remain after the lockdowns lift. Many enterprise XR use cases promise two main benefits: efficiency and cost reduction. Remote assistance can save on costly expert travel to maintenance sites, while reducing operational downtime as well. XR training has proven to be more effective than traditional training, with longer retention and recall for XR trained users, while often taking less time than existing training protocols.

Enterprise Use Case Examples

Remote Expertise	Assembly	Training	Field Service
Huge downtime costs avoided	Reduced error rate and efficiency with complex machinery	Increased retention and recall	Increase first time fix rate and reduce overall downtime
Travel costs eliminated	Integrated verification can quicken completion and increase first time accuracy	Real-time, visual guidance possible on-site	Hands-free capability improves worker capability and safety

While the benefits of XR are clear, it is not always a simple path to deployment. XR HMDs can be costly compared to already implemented devices like smartphones and tablets. These devices also require unique content to be most impactful. Three-Dimensional (3D) assets leveraged in XR need to be created if not available, and even when they are available, they often need additional optimization and development time to properly deploy to headsets. XR is also best when integrated with other enterprise systems, such as the Internet of Things (IoT), Product Lifecycle Management (PLM), Customer Relationship Management (CRM), cloud, etc. Connectivity is especially of concern, as there is immense value in a fully connected XR solution. Wi-Fi is the ubiquitous standard, but cellular connectivity is increasingly being explored for XR applications where Wi-Fi is insufficient or unavailable.

Market-wide standards can help alleviate these pain points and create a more cohesive XR system for adopters. System Integrators (SIs) and platform partners have also quickly become proficient in XR integration and streamlining deployments—combined with broader standards support, XR devices can become as simple to deploy as other, more familiar devices.

Consumer Use Case Examples

For consumers, the conversation is less around integration and Return on Investment (ROI), and more about valuable content delivered to attainable devices. VR gaming is easiest to understand in comparison to traditional gaming—the level of immersion and interactivity possible with VR surpasses anything possible in traditional games. Outside of gaming, content differentiation is less clear, although still present. Volumetric video—a type of 3D video that enables spatial movement within the video—promises a similar level of immersion to VR gaming for other content types from social media and communication to high budget film and TV.

Outside of pure entertainment, education has incredible potential for XR usage. In situations where interaction and visualization deliver significant learning and retention improvement, XR fits perfectly. Serving multi-user content that may be seen in a classroom is a unique content challenge on its own, along with budget restrictions and entrenched content sources slowing potential adoption. Streamlining the creation and delivery of these content types is a boon for all parts of the market. Supporting these use cases in a way that enables flexibility comes down to a combination of connectivity and XR hardware. With mobility a focus for some XR use cases, especially for AR smart glasses, tapping into the existing robust cellular infrastructure is a must. 5G connectivity, either through onboard cellular chips for standalone devices or through tethering to a cellular-connected device like a smartphone, enables users to maintain connectivity despite being on the move. More mobile use cases warrant comfortable and lightweight HMDs, which smart glasses again serve well.

OPPORTUNITIES IN STANDARDIZATION

With a consistent stream of new XR hardware, software, and services hitting the market at a rapid speed, it is easy to see fragmentation creep in. With a need for interoperability across devices and platforms (to ensure a healthy user base and foundation for growth), standards play a pivotal role in XR growth and ultimate potential.

In the traditional media space, the Motion Picture Experts Group (MPEG) and The 3rd Generation Partnership Project (3GPP) have been two of the most prominent names in content delivery standards. Mobile networks are dominant in terms of global data usage, and smartphones are responsible for a substantial amount of total data usage, including video.

Capture, creation, and distribution of XR content all present unique challenges compared to Two-Dimensional (2D) content. For example, a streaming VR use case with interactive content is significantly different from a regular video. Traditional streaming video uses a buffer to allow for changes in network performance and bitrate without interrupting content, but XR interactivity prohibits that buffer, possibly from multiple sensors, meaning a network's uplink and downlink reliability becomes paramount. The nature of VR headsets demands high resolution, high framerate, stereoscopic or multi-view content, and spatial audio—combine these with a need for low latency and high reliability, and the network demand becomes significant. Rather than 25 Megabits per Second (Mbps) for a 4K video with a buffer, bitrate can be closer to 100 Mbps all live and at low latency. Compression and streamlined delivery protocols can help immensely, as can standards enabling interoperability and similar experiences across devices.

MPEG's incumbency in the video space is growing into and alongside XR with a set of standards. The first of them is the Visual Volumetric Video-based Coding (V3C) standard, consisting of both MPEG Immersive Video (MIV) and Video-based Point Cloud Compression (V-PCC). This standard sets up definitions for encoding and streaming volumetric content, which is expected to be one of the most challenging content types for XR in terms of capture, processing, and delivery. These standards provide a 3D/immersive processing layer above traditional 2D video codecs, which means they can benefit from existing hardware-implemented codecs to fast track their deployments, such as Advanced Video Coding (AVC) and High Efficiency Video Coding (HEVC), and will get the additional compression gain of AV1 and VVC in the future. MPEG is active in other XR areas as well, with dynamic mesh coding, haptic, and scene description efforts. MPEG Video Decoding Interface (VDI) enables multiple instances of hardware decoders, a boon for immersive codecs and content.

At Augmented World Expo (AWE) 2023, InterDigital, in partnership with Philips, showcased the first implementation of the V3C immersive video decoder platform, streamlining content creation pipelines for XR content. To make this work, a combination of point cloud compression (V-PCC) unique to volumetric objects, MIV, and Philips' 6 Degrees of Freedom (6DoF) proficiency is required. Use cases for this are varied, with Philips showcasing both live volumetric telepresence and sports content.

InterDigital is a leading developer of wireless communication and computing technologies. Much of its technology appears in 5G devices, networks, and services worldwide and is based on work that takes place in standards development organizations, such as 3GPP. There is a significant amount of activity underway in 3GPP Working Groups (WGs) that focuses on how to enhance the 5G system to better support XR traffic. Most of the early enhancements will focus on how the 5G network can become more "XR aware" and take advantage of this awareness of XR traffic.

For example, enhancements in Release 18 of 3GPP's standard will include the addition of network Application Programming Interfaces (APIs) that allow XR application servers to provide the network with information about XR data streams. By providing the network with information, such as the format of data streams, delay requirements, and relationships between data streams, the network can make more intelligent decisions about how to prioritize packets that carry XR traffic. Knowing how to prioritize, or schedule, XR traffic relative to other traffic that passes through the 5G network is key to ensuring that the user's quality of experience is maintained.

It is well known that XR traffic is often characterized as being "bursty." For example, data packets that are used to construct video frames are often sent in bursts and sent with well-known periodicity. Each burst often represents the data needed to construct a piece of application data (e.g., a video frame), so there is also a strong dependency between the packets within a single burst. 3GPP is making use of network APIs to allow application servers to provide the network with the information necessary to detect these bursts of data. In 3GPP's standard, these bursts of data are called "PDU Sets." The network's ability to detect PDU Sets is a key enabler in making the 5G network more "XR aware."

5G-Advanced introduces awareness of XR traffic requirements in radio protocols, as well as resource allocation techniques where the relative importance of each data packet to the end user experience is considered part of the wireless data transfer. These techniques improve the ability of wireless links to adapt to varying qualities of a radio channel, while accommodating the heterogeneous traffic characteristics of XR traffic. Latency, bitrate, jitter, and packet sizes can all be accounted for.

The 3GPP community will continue making enhancements to the 5G system in 3GPP's next release, Release 19, and beyond into 6G releases. For Release 19, companies that participate in 3GPP are discussing how the network can be XR aware when the XR traffic is encrypted, how the 5G network can compensate for network delay that is incurred by the traffic when the traffic is routed over the Internet, how to deal with highly mobile devices, and how the network can better support traffic that is associated with a device that is tethered (e.g., via Bluetooth) to the 5G mobile device.

OTHER STANDARDS AND XR SUPPORTERS

There are several key enabling technologies that can improve XR experiences over wireless links:

- ***5G Radio Access and Core Networks:*** *Low latency, high bandwidth; better network uniformity (e.g., cell handoff) once infrastructure is in place; required for applications where Wi-Fi is unavailable or unreliable.*

- ***Edge Stream and Compute:*** *Offloading processing at low latency brings benefits in hardware form factor, price, and user comfort.*

- ***Cloud:*** *Data/analytics and workflows are already moving to the cloud and companies that offer support services and platforms are increasingly moving from on-premises and hardware-based solutions to Software-as-a-Service (SaaS) or Platform-as-a-Service (PaaS) alternatives.*

- ***Metaverse:*** *Horizontal support of content across ecosystems; needs solid connectivity and content/service interconnects.*

- ***Codecs and Compression:*** *Maintaining user experience by reducing network requirements; substantial bandwidth needs for high-end XR can be lessened with XR-tailored codecs.*

- ***Development Tools and Platforms:*** *OpenXR and WebXR, Universal Scene Description (USD), and Graphics Library Transmission Format (glTF). Scene descriptors better enable developers to create and distribute object-based/spatial content, while OpenXR and WebXR enable easier XR content viewing for users across devices.*

The metaverse represents the newest and most uncertain environment going forward, along with the question of which standards will be foundational for its services to evolve rapidly. To meet that goal, standards bodies are quickly scaling their own efforts to keep up with metaverse potential. In addition, the Metaverse Standards Forum includes the likes of IEEE, Khronos Group, Open AR Cloud, Open Geospatial Consortium, MPEG, Spatial Web Foundation, and the World Wide Web Consortium. The breadth and depth of this roster is indicative of how important standardization is in a market like the metaverse and, by association, XR. Every element of the metaverse has representation and research in the forum, including the IoT and digital twins, financial technology, consumer content creation, avatars, geospatial data, and more. XR, of course, is included here as well, as a user-facing and interactive element across all these segments.

FUTURE XR CONSIDERATIONS

Despite an already rapid rate of advancement and maturation, several technologies within and tangential to the XR space are still nascent, with many areas left to explore.

- ***Networks will remain at the heart of content delivery.*** *5G is designed with ultra-reliable and low latency services in mind. 5G-Advanced adds XR-specific improvements like XR traffic awareness, XR content identification, and support for multi-user instance, and 6G is likely to continue to support even further the entire set of technologies key to XR use cases. AR and VR have been grouped with 5G for years, as the understanding of AR/VR experiences has revolved around low latency content delivery and reliability. As content grows in scale—both in content scope and actual file size requirements—the need for XR-friendly codecs grows. Low latency applications, such as multiuser instances or edge compute and streaming scenarios, can be catered to separately from other applications. Real-time adjustments, such as resolution and framerate of content, can be made more quickly and accurately to ensure quality of service, from which MR experiences combining real and virtual input/output benefit.*

- ***XR accessories and supporting hardware will be an increasing discussion point.*** *The high frame rates and resolutions provide truly immersive experiences that translate into high bit rates. Generating content at such bit rates creates form factor challenges when powerful XR compute needs to be accommodated in the XR device. This manifests in battery size limitations, heat dissipation, and device weight. In some cases, it may not be possible to host all XR compute in the XR device, resulting in rendering being split across the network. While split rendering* *alleviates the processing requirements on the XR device, it imposes additional delay constraints to generate a truly immersive experience (e.g., motion-to-photon latency < 20 Milliseconds (ms)). Failure to meet the delay requirements can result in motion sickness and poor user experience. Haptics, gestures, voice, eye tracking, and traditional controllers all have room for improvement—one area being accuracy. The greater the accuracy, often the greater the data rate and thus the need for an efficient processing pipeline.*

- ***While some of the hype has died down from its peak, the metaverse remains a staple of the future of XR.*** *Given ABI Research's view of the metaverse, mostly as an interconnectivity back end for varying technologies going forward (XR is one of many), metaverse and standards efforts go together perfectly. Identifying technology headwinds across numerous platform and hardware types is a significant effort, with fragmentation a constant threat of hampering potential and growth. Standards organizations working in XR, metaverse, and other technologies can avoid those headwinds from the beginning.*

- ***Volumetric media, as mentioned earlier, presents a combination of incredible content and incredible content challenges.*** *Volumetric video literally adds another dimension to video content, and with that comes unique compression and streaming challenges. Rather than a single "view" for traditional video, volumetric video holds data for numerous parts of the scene—the data rate for uncompressed high-fidelity volumetric video can be orders of magnitude higher than even 3D video, so it requires unique compression capabilities to get file sizes to a more manageable level. Integration of all media, such as haptics and volumetric audio, in a scene representation and composition is also a key challenge to achieve in real time, low latency, and constrained devices.*

- ***Multiuser support for XR presents unique challenge through both the type of content and the level of fidelity for that content,*** *with all users needing to synchronize their specific virtual environments at low latencies. Add to this the high average resolution and framerate required for XR, and network nodes with even just a few users can be overwhelmed. Better parallelization for XR, alongside specific standards-derived feature sets as seen with 5G-Advanced, will keep progress brisk for multiuser experiences.*

- ***Positioning and user tracking remains an important element of XR.*** *Head tracking has grown from outside-in tracking (requiring an external tracking station), to inside-out tracking where the headset can position and track itself independently. However, there is interest in expanding tracking outside of just physical hardware being worn by the user. Ultra-Wideband (UWB) is being researched as a broad positioning method to supplement other XR tracking, especially in relation to other UWB devices. Some forward-looking efforts are expanding beyond that and using radio for sensing applications, instead of traditional cameras and tracking sensors.*

While there is significant progress to be made in every facet of XR, there is already a very strong foundation on which to build, despite only being viable within the last 5 to 7 years. With so many moving parts, fragmentation can quickly stall innovation. Standards can navigate around those stalls and ensure compatibility across systems that users find valuable. Improvements in content capture, encoding, and distribution specific to XR will combine with hardware-level improvements to spur the market from all sides—hardware, software, services, and networks.

Published August 2023
ABI Research
157 Columbus Avenue
New York, NY 10023
Tel: +1 516-624-2500
www.abiresearch.com

ABOUT INTERDIGITAL

InterDigital develops mobile and video technologies that are at the core of devices, networks and services worldwide. We solve many of the industry's most critical and complex technical challenges, inventing solutions for more efficient broadband networks, better video delivery, and richer multimedia experiences years ahead of market deployment. InterDigital has licenses and strategic relationships with many of the world's leading technology companies. Founded in 1972, InterDigital is listed on NASDAQ. InterDigital is a registered trademark of InterDigital, Inc.

ABOUT ABI RESEARCH

ABI Research is a global technology intelligence firm delivering actionable research and strategic guidance to technology leaders, innovators, and decision makers around the world. Our research focuses on the transformative technologies that are dramatically reshaping industries, economies, and workforces today.