TRAILBLAZING WITH 8K LIVE VR AT BEIJING 2022

ABSTRACT

Prior services with 4K VR highlighted the potential of live VR video services but fell short of the resolution required for a great live sports experience. 4K resolution is not nearly enough to meet the expectations of viewers typically watching HD television. Intel is answering business needs with end-to-end solutions powered by Intel® Xeon® processors that are designed to deliver live 8K VR and 360-degree video experiences to billions of viewers across the globe. One such solution was on display at the Olympic Winter Games Beijing 2022, where Intel VR technology platforms enabled the capture and production of the first-ever fully produced live 8K VR video broadcast for Olympic Broadcasting Services (OBS), which then made it available to rights holding broadcasters (RHBs) worldwide.

Intel is a Worldwide Olympic partner (TOP) for delivering VR technology platforms to the Olympics and Paralympics. Using the latest processors, graphics cards, accelerators, cameras and distribution technologies, Intel collaborated with OBS and other partners to produce and deliver a trailblazing 8K live VR experience. In an industry first, this collaboration streamed over 100 hours of fully produced Olympic events in Beijing in VR, with graphics and commentary. Live coverage included six sports including hockey, snowboarding and figure skating in ground-breaking 8K resolution for both VR180 and VR360. The user reviews were excellent, and we believe that a VR production at this quality meets the bar for attracting a wide, general audience.

INTRODUCTION

VR video, live and cinematic, experienced a hype cycle rollercoaster over the last 7 years. The expectations were severely inflated around 2016 and there was a deep trough of disillusionment around the turn of the decennium. Throughout this hype cycle, technology development continued: VR cameras, headsets, and distribution technology have all steadily improved. Technology advancements and industry shifts heralded a renewed interest in VR and this time around, the technology exists to meet the quality demands for a compelling, sticky user experience.

While prior live sports trials and services proved their potential, the 4K 360-degree resolution fell short of what is required for a great VR experience. VR video that is spread across 180 or 360 degrees requires a very high source resolution to meet the expectations of viewers typically watching HD television. With the latest cameras, graphics cards and distribution technologies, we can now produce and deliver a trailblazing 8K live VR experience. In an industry first, we partnered with OBS who streamed over 100 hours of Olympic events in Beijing at ground-breaking 8K quality for VR360 and VR180. Enabled by Intel and various third parties, OBS produced and distributed events in six sports: ice hockey, snowboarding, free style skiing, curling, short track, and figure skating as well as the opening and closing ceremonies. All events were streamed live, across the globe, to users in the USA, China, as well as to select users in Europe.

The events were distributed to headsets, mobile devices and web players, and the content included live events as well as on-demand features. This paper's focus is on the workflow for live events experienced in a VR headset.

The production used five VR180 cameras in a director switched "VR Cast" which the user experienced in a virtual suite environment. They also had access to one VR360 camera for total immersion. The user could freely switch between the 180 and the 360 feeds. Events were available for full replay immediately after the event, with the app also showing event highlights and special features – all in VR. As a part of the experience, a virtual HD "Jumbotron" showed the TV broadcast, with broadcast announcers in the audio feed.

The content was available in the USA for users of "NBC Olympics VR by Xfinity" app, and in China with CCTV's Yangshipin app. Chinese users also had access to web streams in mobile apps and browsers. App store reviews raved about the significant quality increase over similar VR experiences in the past.

WORKFLOW

We will first describe the end-to-end workflow and then zoom in on a few details that are interesting to the IBC audience. This is the high-level workflow:

- *Capture & Transmission:* camera choice, camera placement and transmission
- *Production:* receiving feeds, multi-camera production editing, adding graphics and a multi-channel audio mix, generating mezzanine feeds
- *Contribution:* sending the mezzanine feeds for consumer transcoding
- *Cloud Services & Transcoding:* producing streams suitable for distribution to consumer devices
- *Distribution:* the distribution to consumer devices over CDNs
- *End-user consumption:* the consumer player and the interactive environment

We describe each of these elements below.

Capture & Transmission

Intel worked with a production and immersive content creation partner (Cosm) which was responsible for the overall solution and provided both production and client-side apps. The most significant technical change with respect to past workflows is the capture in 8K. The 360 feeds were captured in an equirectangular projection ("ERP") of 7 680 x 3 840 resolution. Two different types of 8K cameras were used, one smaller model at 30fps and a larger camera at 60fps for the best quality. The smaller camera was used in locations where space was at a premium, like between two curling sheets. The larger camera was used for the Opening and Closing Ceremonies.

While there are a few 8K VR360 services, even live (some sport apps offer up to five 8K VR360 feeds), we are not aware of any live 8K VR360, at 50 or 60 fps, services for VR headsets. Even if they do exist, Beijing 2022 was certainly the first event covered in 8K *VR180*, a major quality step over 8K VR360. All VR 180 feeds were captured at 60fps. A framerate of 60 or 50 fps is the minimum requirement for a good sports viewing experience, yet many VR productions still use 30 or 29.97 fps.

Five camera bodies fitted with 8-15mm f/4L Fisheye USM lenses captured the VR180 footage at 7 680 x 4 320 resolution. The cameras provided 17-stop dynamic range – well into HDR range. We collaborated closely with the body manufacturer to enable the camera to stream and 8K RAW feed over IP from the camera to the Camera Control Units (CCUs) located at the International Broadcasting Centre in Beijing (further referred to as “Beijing IBC"). Each camera feed was compressed to around 2.5 Gbps using wavelet-based compression, transmitted over IP using SRT and then decompressed, delayered and processed by the CCU. As the cameras used typically produce raw sensor readings meant for cinema post-production, we also worked with the vendor to add spatial denoising to the CCU and to enable full control of camera and lens from the Beijing IBC. A suite of camera shading options was also added. The CCUs outputted the streams in uncompressed 8K 59.94 over Quad SDI in 2 Sample Interleave.

The difference between 8K VR180 and 8K VR360 is significant. This is because in the VR180 feeds, the full 8K resolution was used for a (180-degree) hemisphere instead of a full (360-degree) sphere, more than doubling pixel density over a traditional 8K VR360 video. It is *more* than double because the 180 feed didn’t cover a full 180-degree arc vertically.

The ERP-equivalent resolution is a good metric for the quality. For the VR180 feed, with the zoom lens set at 13 mm, it was slightly more than 15K. This means that in the heart of the fisheye image, the resolution is about 15 thousand pixels per 360 degrees. We calculated this by measuring the average horizontal resolution in pixels per degree over the entire sensor, where the horizontal angle captured by the lens was 184°. This gives an ERP-equivalent resolution of 7680 / 184° x 360° = 15026 pixels on the equator. This equates to 41.7 pixels per degree, on both the horizontal and vertical axes. This is an average: on the horizontal axis, the fisheye lens gives a resolution in the centre of the image that is approximately 15% higher than at the edges. Note that at the 13mm lens setting, the full sensor was lit. The consumer transcode was done at a lower resolution; more on this below.

We note that this resolution is significantly higher than the ERP-equivalent screen resolution of the devices that viewers used. That screen resolution is about 8K for the headset models commonly used in the USA and China.) Extra resolution is still very meaningful to counter loss of resolution in the various transformations: from fisheye to cubemap for the distribution format, and then the dynamic mapping from video raster to the display as users look around to follow the action. See Section “Transcoding” below for details on the distribution resolution.

Figure 1 - Camera to Headset

As noted above, we used five VR180 cameras for the production, with a director producing a single VR180 feed by switching between these. The cameras were placed as close as allowed to the action, in such a way that collectively, they could cover all of it. For example:

4 cameras along the halfpipe and one at the end, and the same set-up for the moguls. For ice hockey, one camera at the main TV platform (in the stands), two behind the goals and another two in the corners. As it is impossible to zoom in, being close to the athletes is paramount to a great VR experience – like it would be for a “real” visitor.

With cameras as close to the action as possible, and given the harsh weather of winter sports, a lot of attention was paid to weather-proofing and camouflaging the cameras. We created custom 3D printed lens visors to protect them from snow and rain (the default visor would be in the field of view of the lens), we added a lens dew heater, and combined all power and IP conversion in a weather-proofed field kit. Lastly, we had custom covers made, in white (for outdoor use) and black (for indoor use). The small size of the camera and the careful attention to these details were instrumental in securing great camera positions for our VR cameras, which still took careful planning and coordination with Olympic Broadcasting Services (OBS).

The TV compound setup was compact, consisting of IP equipment powered by Intel Tofino switches that sent the camera feeds back to the IBC. The Freestyle skiing and snowboarding events took place in Zhangjiakou, 180 km from the Beijing IBC. The feeds for these events were transmitted through an extra hop at ZBC (Zhangjiakou Broadcast Centre) before going to the Beijing IBC over an L2 connection. An NDI signal was sent back from the IBC to the Venue Compound for monitoring purposes. The switches allowed us to relay the 2.5 Gbps traffic from the camera feeds.

Production Enablement

The VR-Cast was produced in the Beijing IBC. A director, assisted by a technical director (TD), mixed the five incoming VR180 feeds into a single output stream using an 8K switcher. We also placed a virtual jumbotron in 3D space of the VR180 scene, as well as additional graphical elements such as score and time. The virtual jumbotron, containing the TV program feed, was inserted as an HD feed. It would normally sit at an unobtrusive yet visible location, usually above the action. The virtual jumbotron, much like its in-stadium counterpart, provided the close-ups and replays that viewers didn't get from a VR camera. During replays when there was no live action, the director enlarged the virtual jumbotron and brought it forward to be front and centre, using a simple yet pleasing animation.

Care was taken to have a reasonable amount of scene cuts across the VR180 cameras, following the action. While it had been believed that scene cuts had to fade through black to prevent user disorientation, this proved unnecessary in practice. VR headset users have no problem with hard cuts at all. Camera motion is much more disconcerting for VR users, but all live cameras were fixed, without motion, which provides the best user comfort and allows for easiest orientation. That said, camera cuts can still be a little confusing if they move the focus of the action to a completely different location in the VR180 sphere. That's because the user needs to quickly find that new location by turning their head and looking around. The director tried to ensure that camera switches would not have the subject jump out of view of the user. For the Halfpipe, this implied that crossing the line was actually helpful; it allowed the viewer to pick up the athlete more intuitively after a cut.

Sometimes it was inevitable to do cuts that moved the action for the viewer, such as for the Mogul Run, where four cameras were placed along the slope and the action switched from bottom-left to top-right after a cut. When this was the case, the director ensured that the cuts were predictable across runs, so that users would know what to expect and how to move their head during a run. Anecdotal evidence shows that this worked well.

The VR360 feed had a virtual Jumbotron as well; it was a bit larger than in the VR180 feed, as there is more room in a full (360-degree) sphere than in a 180-degree hemisphere.

Real-Time Graphics

Another partner provided the graphics production for the VR180 and VR360 feeds. All graphics were generated in a game rendering engine, to allow placing them in 3D space. We added graphics to the 8K feeds in real-time, using a dedicated operator.

They were processed in the native projection of the camera/lens combination. Graphics processing for the VR180 feed happened in the fisheye domain; the graphics where distorted to match the fisheye equisolid projection. We did this to minimise the number of reprojections in the signal, as all such projections inevitably come at the cost of some loss of fidelity. The cloud transcoder did the conversion to the cubemap domain; more on that below. The graphics were generated as key-and-fill and sent to the 8k switcher. The use of the switcher APIs allowed us to only use 2 graphics servers by automatically setting the scene for the camera in preview and alternating the graphics servers between preview and program.

Intel also collaborated with a third party to upgrade the firmware of the router to support 8K in 2SI. The router enabled us to route all feeds as needed for the production and enabled audio embedding of OBS audio with Rights-Holding Broadcaster (RHB) commentary.

To allow real-time insertion of the jumbotron (containing the HD broadcast feed) and other graphics elements such as score and time into the 8K p60 feeds, we needed to create our own rendering pipeline. It used multiple virtual cameras that rendered a static point of view (PoV) and projected that into a 2D texture buffer. We also created a High-Level Shader Language (HLSL) custom shader that was applied after the 3D environment was rendered. This allowed us to match the graphics with the live, fisheye-distorted footage.

The five different VR180 cameras each provided their own unique view of the event and required insertion of the graphics elements at different positions in the image. This was necessary to ensure these graphics did not sit on top of the action, which might be in a different location for each of the cameras.

We chose an upstream approach in which the rendering engine sends a fill/key signal to the switcher. The graphics controller listened to the switcher and would know what camera was currently loaded in preview. The rendering engine signaled the appropriate coordinates for each graphics element to the game engine. Once the graphics elements were correctly positioned, the technical director could take the shot live. We employed two individual game rendering engines that alternated between live program and preview. This allowed the technical director to call up any shot on one engine while the other engine was on air. We used a local server that listened to the switcher events and published these on a centralized real-time database with the controller picking these up. This ensured synchronization between all engines and that controller.

The controller also sent additional data to the rendering engine such as titles, locations of objects, and sizes and it triggered animations. The operators used it to create and drive animations during the live events, such as the jumbotron coming forward and becoming larger when there was a replay in the program feed.

Creating Social Media Feeds

We created two separate output feeds for use on various websites and on mobile devices, and we needed to give it the same look-and-feel as in the headset. Using a third party solution, we created an After Effects template that took the 180 feed, projected it onto a 360 sphere, and then overlayed the virtual suite and the other graphics elements on top. The result was a VR 360 ERP video used for social media. Since the virtual lounge was customized for the individual Rights-Holding Broadcasters, we had to create these social media feeds separately for each of them. The three versions were for NBC, CCTV and OBS.

We also inserted the proper audio channels: English for NBC and OBS, and neutral audio for CCTV. (Note that the headset feeds contained all audio tracks, with the application selecting and decoding the correct one. The social media feeds only had one stereo track.) We then inserted the metadata required to recognize the feeds and clips as VR videos, egressed them to the different RHB’s servers where they could take these feeds and clips for publication on social media channels.

Contribution

We now move back to the tiled workflow. The contribution was straightforward: the VR Cast VR180 and VR360 feeds were encoded at 120 Mbit/s in HEVC, and then transported from the Beijing IBC to a public cloud data centre in Tokyo using Secure Reliable Transport (SRT). We employed 8K encoders running on Intel processors for the real-time mezzanine encode. The transport was done over a dedicated fibre connection to Tokyo and this line was cross-connected at the Tokyo data centre. We used an SRT broadcaster, so that different parties could take the SRT streams: one party took the feeds for distribution transcoding, another had select LED domes pick up the streams for display, and Rights Holders also used these streams, e.g., as a live studio backdrop. See Figure 2 for an example from Tokyo where similar backdrops were used.

Figure 2 - Studio Backdrop using an Immersive Camera as used for Tokyo 2020

Transcoding

We then transcoded both VR feeds into tiled and non-tiled adaptive bitrate ladders. The non-tiled (HLS) feeds were mainly consumed in China, where millions of viewers accessed the live VR content on CCTV websites and on mobile devices. The tiled feeds were used in headset applications.

In VR applications, the user only sees a part of the image: about $1/8^{th}$ for a 360 feed, and $1/4^{th}$ for a full hemisphere for VR180. This means that legacy, non-tiled streaming systems transmit many more pixels than required for the viewer to see the action. The same applies to the decoder: it decodes pixels that the user will never see, limiting the resolution for the video that the user *does* see.

Tiled streaming overcomes this by cutting the image in tiles, and then only streaming the tiles that are in view. Since the client application knows where the viewer is looking, it can determine which tiles are required and request those from the CDN. The client-side application then reassembles these bitstream snippets into a single legal bitstream and sends that to the hardware HEVC decoder in the device. The decoded frame looks like a sliding puzzle, and the tiled video player takes the individual tiles and projects them in the right place in user's view. A few tiles are used to provide a lower-resolution version of the video that covers the full sphere (360) or hemisphere (180). This prevents users from seeing black patches when they move their head. If network speed is sufficient, this background virtually unnoticeable. See Figure 3 for graphical depiction of tiling (note that for illustration purposes we show an ERP instead of the cubemap that is used in the actual transcode).

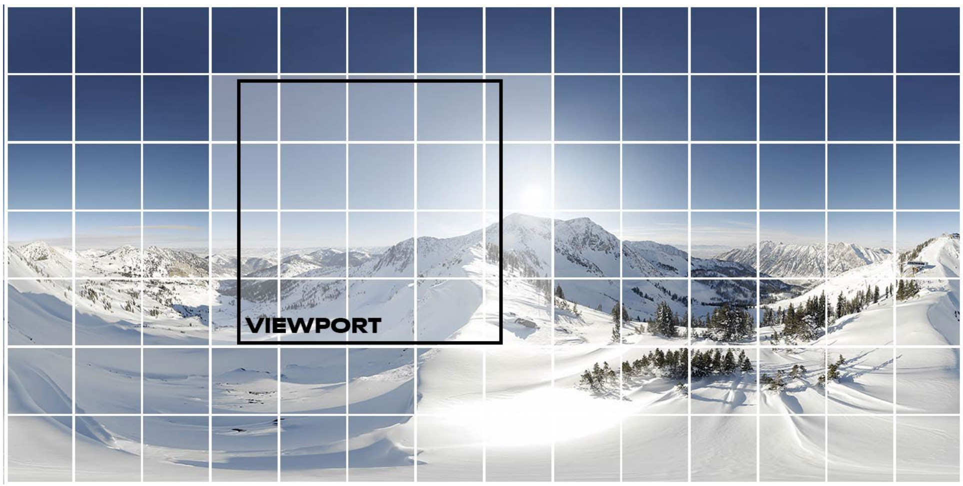

Figure 3 – A Tiled ERP

Tiled streaming is inherently robust, because the decoder can determine, just before a frame needs to be sent to the decoder, which tiles have arrived and which have not. If a tile is missing, the decoder just leaves it out of the frame. The system is also inherently scalable, limited only by CDN capacity. All intelligence to determine which tiles should be retrieved for the best user experience resides client-side, and tile retrieval consists of simple http requests, without per-user processing. Our tiled video player is aware of the viewport, decoder capacity, network conditions and other parameters, and uses these parameters to optimise the experience.

Our system uses the HEVC coder with its tiling features, and it is based on the *Advanced Tiling Profile* of MPEG's *Omnidirectional Media Format* (OMAF) standard 2. The 8K VR360 video used about 100 high-resolution tiles, and the VR180 feed about 90. Note that the tiles are not necessarily square; our system calculates the optimal tiling configuration.

Before tiling, we converted the ERP into a cubemap, which is a more efficient representation of the spherical video as an ERP is very wasteful at its zenith and nadir. The 180 fisheye image is converted into half a cubemap. To allow adaptive bitrate streaming, three cubemaps were created at different resolutions to provide 3 ABR levels. Beijing 2022 marked the first time that such an adaptive bitrate tiled streaming was used, both for live and on-demand content.

To allow for the best user experience, we do not only provide 3 ABR levels, we also encode three different Group of Picture (GOP) versions for each. This allows the client to respond immediately to head motion (using the shortest GOP) and to be as efficient as possible (using the longest GOP). The client is opportunistic in switching as fast as possible to the most efficient GOP version for any individual tile. Our metrics show that 80% of such tile switches complete in under 50 milliseconds; and 98% of tile switches completed in under 200 milliseconds.

Running in a Tokyo public cloud data centre, our tiled transcoder takes the following steps. First, using GPUs, it reprojects the VR video into a cubemap, creating different resolution versions at the same time. After determining the best tiling scheme for each resolution (this happens only once) it then extracts the tiles for each of the versions, with tiles having a small overlap for seamless reconstruction. All tiles (small videos) are then encoded in the three different GOP versions. Our system uses a manager task that dispatches many parallel encodes as well as packaging jobs. Tiles are packaged together into ISO files in a way that optimises CDN cache hits. We use a modified open-source SVT-HEVC software encoder provided by one of us (Intel), running on many thousands of CPU cores working in parallel for the live workflow. The encoder modification constrains motion vectors to allow undistorted tiled decoding.

While the global 8K distribution was already a first, the quality of the VR180 feed was unprecedented, as we noted above, with a source signal that had an360 ERP-equivalent of about 15K in the centre of the sensor. The VR headsets employed for Beijing 2022 have about an 8K-equivalent in the viewport, but oversampling is useful, and the quality increase of distributing in an ERP-equivalent of 10K - 11K as compared to 8K is clearly visible. This is because the path from camera to headset inevitably involves a number of projections: from fisheye or ERP to cubemap (for efficiency) and then from cubemap screen pixels on the display. (The VR360 has another step: it takes input from multiple sensors and fuses that into a single cubemap). The ERP-equivalent resolution of the distributed VR180 was about 12.5K, 11K and 8.5K respectively, denoted as "High", "Medium and "Low" in the app. For the VR360 feed, these resolutions were 8K, 6K and 4K.

Professional viewers noted that, for the VR180 feed, the difference between the "High" and "Medium" levels was small, if noticeable at all, but the different between "Low" and "Medium"/"High" was very clear.

The non-tiled feeds were traditional Adaptive BitRate (ABR) HLS feeds. They were AVC-encoded for use in web players, at 4 levels from 4K down to 1080p still represented as a fisheye (VR180) and an ERP (VR360) respectively. The total bitrate for all tiles across all ABR levels in all GOP versions – required to give the best user experience – was about 700 Mbit/s per feed, with an additional 45 Mbit/s for the HLS ladder.

After a live event ended, we created on-demand manifest files for all feeds. This provided app users with full replays of all live events, immediately after they ended.

Distribution

After transcoding, we egressed the feeds to a dedicated origin server in Tokyo. Using a fast and reliable replication protocol, the content of this origin server was replicated to additional origin servers back into China as well as the U.S. West Coast, the U.S. East Coast, and Europe, in that order. We used two CDN providers, one in China and one for the rest of the world. The CDNs were configured to pull from our dedicated origin servers.

Providing bitrates for tiled streaming is harder than for regular (ABR) streaming, for three reasons. First, bitrates for tiled streaming depend on user behaviour. Frequent head motion results in more tile requests and for, on average, less efficient GOP representations. Second, tiled streaming is "greedy": it will use the available bandwidth to quickly download tiles, which will cause the bitrate to spike. This is part of the inherent robustness of tiled streaming: when bandwidth is (temporarily) limited, the system will just respond a bit slower in updating tiles to high resolution. Third and last, bitrate can depend on the content and notably the amount of motion in it. While providing bitrate information is very informative in this context, this implies we can only give rough, estimated, average bitrates.

At the highest level for the VR180 video, clients consumed 25-35 Mbit/s for the 180 feed, and 12-18 Mbit/s for the 360. The lowest level centred around 15 Mbit/ for the VR180 and 6 Mbit/s for the VR360.

Consumption

Intel engaged Cosm who developed the VR headset app that was white-labelled and customized for Rights Holding Broadcasters NBC Universal and CCTV. The app incorporated a tiled video player software development kit (SDK). Users had access to the VR headset app through a popular app store in the USA and another app store in China.

During events and replays, users could choose between the 360 and 180 feeds. The VR180 feed was presented in an attractive virtual suite, decorated with Olympics posters, graphics and gadgets. The suite had a large virtual window, which made the user feel like they were looking out to the action, and often – depending on camera position – they were very close to it. In ice hockey, for instance, two cameras were positioned against the glass wall behind each goal, and players would often bodycheck into that wall and seemingly into the user. Many viewers will have jumped back in their seats when that happened – that's how real it felt. The virtual suite was designed such that it covered the edges of the video on all four sides, enhancing the sense of presence. Figure 4 gives an impression of the environment; users only see a small part of this in their headset – about 12% of the full ERP.

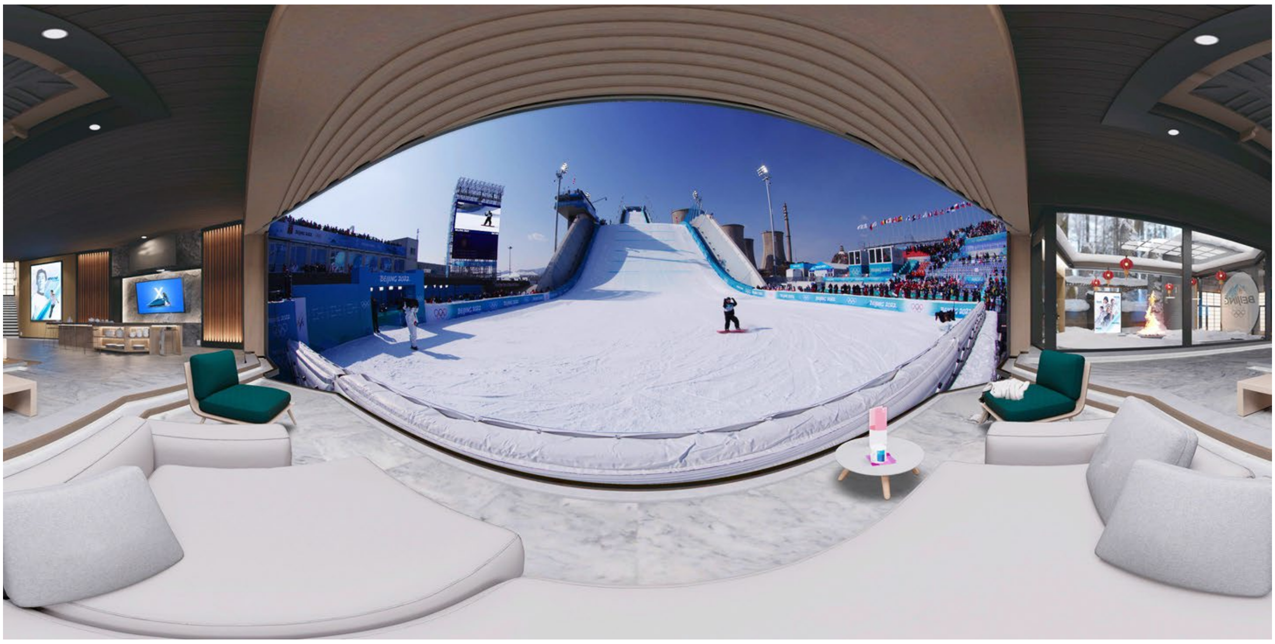

Figure 4 – VR Headset App Experience

The NBC app gave users the option to talk with their friends, in real-time, when watching the live events. Using a webRTC peer-to-peer voice communication module, sports fans were able to set up "Watch parties" with up to 4 friends. While the glass-to-glass latency for the

VR feeds was about 80 seconds, all viewers had that same delay, and the voice communication was instant. As a result, fans could still freely talk about the events they were watching and cheer for goals, without spoiling the experience for the other participants in the watch party.

User Reviews from *NBC Olympics VR by Xfinity* App

An Amazing Way to View the Olympics – by **rbuccigrossi** - Feb 5 at 10:39 PM
"I am a large Olympics fan (...) While I have a pretty nice projector viewing setup it pales in comparison with the VR experience: Olympics in the Quest is simply breathtaking. (…) If you are an Olympics fan, this is *AMAZING*. Thank you for providing this experience!"

Amazing. That is all – by **Jegaysus** - Feb 5 at 6:47 PM
"Watching the Olympics like this is absolutely the way to go."

Awesome – by **FuzzyAlien** - Feb 6 at 11:18 PM

"Lovely app. Interface is intuitive and performs well. Very immersive and I love the wide angle views. It provides a completely new perspective on the skills of the athletes and especially the speed!"

Amazing Quality - by Dano7568 - Feb 4 at 1:09 AM
"I watched the Moguls qualifiers and it was almost like sitting in a room and watching the event right there. Which of course is the point of the app! Also the Winter Olympics virtual environment (lobby) is very cool!"

WHAT'S NEXT FOR VR IN SPORTS?

The Beijing 2022 Olympics significantly raised the bar for VR sports distribution in many ways. It was the first event that had a fully produced 8K VR workflow with switching between multiple cameras, dynamic graphics insertion, and social features. It was also the first time 8K VR360 and VR180 were distributed live to a worldwide audience. It was certainly the first experience – live or on-demand – that employed 8K VR*180*, a resolution that significantly, and visibly, improves over VR360. In addition, this was a 60fps workflow, a framerate that a proper sports experience demands.

We believe that at this quality of production and distribution, live VR video has finally crossed the experience threshold that enables a great and sticky experience for a mass audience, and we hope more such events will follow the Beijing 2022 example. Users agree: the reviews they left both app stores were very positive about the quality of the experience.

At the same time, this is certainly not the end of the road for VR video. Better headsets will appear: lighter, and with higher resolution. When they do, the production and distribution workflows described in this paper are ready to improve the user experience even further. At the native 8K sensor resolution, the VR180 signal clearly out-resolves the headset. This means it's now up to the headset makers to take the next step.

Also, VR360 cameras will start to go beyond 8K – they already do for on-demand content, and they will soon for live as well. VR360 can then match the amazing quality of 8K VR180. The tiled distribution of VR360 at, say, 11K resolution will be the same as that of the 11K ERP-equivalent VR180 distribution used for Beijing 2022. It will have the same viewport resolution and, therefore, the same bitrates, and can benefit from the Beijing 2022 distribution workflow.

ACKNOWLEDGMENTS

The authors acknowledge the work of many individuals and companies to enable the Beijing 2022 VR experience.

Producing Partner: Cosm

Graphics processing by Twizted Design

8K mezzanine encode performed by Spin Digital

Tiled transcoding and distribution by Tiledmedia

CONTRIBUTORS TO THIS PAPER

Intel: Courtney Willock, Ravi Velhal

Cosm: Ryan Cole, Fabrice Lorenceau, Devin Poolman

Twizted Design: Jan Umansky

Tiledmedia: Rob Koenen, Marius van Eck

REFERENCES

1. Joint Collaborative Team on Video Coding (JCT-VC), ISO/IEC 23008-2:2013 / ITU H.265 - High Efficiency Video Coding

2. Moving Picture Experts Group, ISO/IEC 23090-2, Omnidirectional Media Format, 2nd Edition

3. Intel, Scalable Video Technology, https://www.intel.com/content/www/us/en/developer/articles/technical/scalable-video-technology.html, last accessed 1 May 2022