Advanced compression and streaming tools for ultra-high definition immersive media with live experimentation

Maciej Głowiak[PSNC] *mac@man.poznan.pl,* Maciej Strozyk[PSNC] *mackostr@man.poznan.pl,* Szymon Malewski[PSNC] *szymonm@man.poznan.pl,* Mauricio Alvarez Mesa[SD] *mauricio@spin-digital.com,* Sergio Sanz-Rodriguez[SD] *sergio@spin-digital.com*

[PSNC] Poznan Supercomputing and Networking Center, [SD] Spin Digital Video Technologies GmbH

Abstract

Advanced next-generation media for immersive spaces require efficient video compression and transmission, which is related to the implementation of modern encoding and streaming methods. The use of prototype devices existing currently on the market requires time and integration, but opens up a huge space for experimentation and the use of research results in practice. In this article we present modern solutions for encoding high-quality and high-resolution media, as well as for obtaining video content and media streaming. We present the results of experiments of long-distance 8K streaming with low latency. Artists, also associated with the STARTS initiative, participated in these works.

Introduction

In order to provide users with the best viewing experience, novel immersive media requires support for an ultra high resolution – 8K and beyond - and, therefore, processing and transmission of huge amounts of data. So customized and efficient data compression algorithms play an extremely important role. This is particularly important if these media need interaction mechanisms and user interaction with the content displayed.

The European Immersify project funded under the H2020 program addresses these challenges in its research. Among the issues addressed in the project are the creation of cutting edge tools for the next generation of immersive media. The work is focused on improving the quality of immersive media using advanced compression technology as well as enabling immersive media for multiple display environments and multiple devices. The Immersify project undertakes to perform real-time demonstrations of next generation immersive media and creation of new immersive content in order to demonstrate its results.

To present the possibilities of the new tools, the project cooperates with artists either as part of the VERTIGO STARTS artistic residencies or in collaboration with independent artists experimenting with high-resolution immersive and interactive content.

High-quality encoding for UHD immersive media

A new generation of immersive media requires a combination of many new aspects and technologies. To support immersive installations such as new cinemas, dome theaters, 8K and 3D large format displays, CAVE systems and deep spaces, in addition to support for high-resolution video in 8K and 16K, it is also necessary to guarantee a high frame rate (120 or even 240 fps) and high image quality by using High Dynamic Range (HDR), Wide Color Gamut (WCG) and 10- or even 12-bit depth. As a result, the uncompressed data stream reaches a very high level. For an 8K x 8K image at 120 Hz it may reach even 212 Gbit/sec, which gives 1.6 TB/min. Processing of such a signal is very demanding, not to mention the network transmission. To support such a signal, compression technologies are used. The Immersify project develops advanced coding tools based on the HEVC/H.265 standard in order to provide the required compression efficiency.

The use of optimized HEVC implementations allows to reduce the 8K data stream to 50-120 Mbps while preserving excellent image quality. As confirmed by QoE tests carried out in the project for various types of content (camera image, CGI) such a bitrate range is sufficient to maintain a quality practically indistinguishable from uncompressed stream, at least for typical cinema productions. As QoE tests have shown, only CGI (e.g. point cloud rendering) requires much higher bitrates.

In the Immersify project there have been also conducted experiments with different real-time 8K encoding solutions:, Advantech hardware encoder (VEGA-3304 card based on Socionext chip), GPU-based NVEnc encoder, and Spin Digital CPU-based encoder (Spin Enc Live), all for 8K or beyond. They all offer real-time encoding of at least 8K 60p video with various bitrates and parameters. Formal subjective experiments were

conducted in the project using Spin Enc Live, showing that broadcast-grade quality can be achieved even at 50 Mbps.

	Advantech VEGA-3304	NVEnc	Spin Enc Live
Type	Double width PCI Express Gen3 x16	2x Nvidia Quadro P5000	Software, CPU-based
Max. resolution and frame rate	7680x4320 at 60 Hz	7680x4320 at 60 Hz	7680x4320 at 60 Hz (or equivalent pixel rate)
Chroma formats	4:2:0, 4:2:2	4:2:0	4:2:0
Bit depths	8-bit, 10-bit	8-bit	8-bit, 10-bit
Maximum bitrate	500 Mbits/s	1300 Mbit/s	250 Mbit/s

Table 1. Specifications of real-time 8K HEVC encoders

Experimental 8K and beyond content production in collaboration with artists

Creating 8K content is extremely demanding. First of all, 8K recording equipment is still in the prototype phase and not easily available on the market. Secondly, post-production of immersive content, storage and processing this huge amount of data is very time consuming. Technically, the workflow for typical 8K film productions is currently feasible, except usually long time required for processing, but creating 8K stereoscopic content requires the use of specialized 3D rig and perfect synchronization of cameras or photo cameras. Photo cameras may be used for timelapse shooting, which means taking thousands of photos and generating videos frame by frame. For this last application, German photographer Martin Heck from Timestorm Films was invited to create amazing images of the scenery of Gran Canaria[1].

PSNC, which has 4 SONY F65 cameras modified to record and stream 8K signals (8K 59.97p), created short video clips that were used as test and promotion content by a number of companies at various technology fairs (IBC, NAB, InterBee), conferences (TNC) or film festivals, just to mention Spin Digital, Leyard, Intel or Advantech. In 2019, the first screening of 8K films (organized by Immersify) took place at the famous Festival de Cannes, where a dedicated movie theater for 8K screening was rented[2]. The 8K video was also presented at the Ars Electronica Festival in their Deep Space large immersive installation and was available for visitors. In this installation, UHD video is displayed on two screens - the front and the floor, each is 16 x 9 meters in size and is operated by several projectors in blending mode.

The second, and even more demanding multimedia content is an immersive 360°/VR video presented not only on Head-Mounted Displays glasses which still offer insufficient resolution, but mainly on large-format displays, such as dome theaters or CAVE systems. Existing 360° cameras on the market, such as the Insta360 Pro, are insufficient in terms of offered resolution or encoding quality, so the PSNC built a 360° rig based on 8 BlackMagic Mini Pro cameras, which are capable of recording 4K images each. However, the post-production and stitching of these panoramic signals take many weeks of work, so creating such advanced content in 16K-20K resolutions is complicated and a special workflow is needed here. To combine images from multiple cameras (this process is often called “stitching”) the PTGui software was used. For further processing such as post-production, color and visual corrections as well as conversion between Equirectangular and CubeMap projections, we used the BlackMagic Fusion 16. This process is time consuming, as the raw 3 minute-long footage from all cameras take 104 GB, but after stitching and conversion to PNG files we obtain 1,85 TB of data. Final HEVC encoded file takes 5-7 GB depending on the bitrate and the quality.

The third group of content created in Immersify is the use of CGI graphics and the generation of high resolution images using computer graphics. As part of the VERTIGO STARTS artistic residency, German artist Theresa

[1] Island in the sky II, Timestorm Films, https://immersify.eu/content-demos/island-in-the-sky-ii/

[2] Immersify at Cannes XR, https://immersify.eu/news/immersify-at-marche-du-film-festival-de-cannes-2019/

Schubert created the work Immersive Minimalism ("Always dead and alive"[3]), in which the development of the cellular-automata, where each virtual cell takes up 1 pixel of the image, is artistically depicted. The second artist was Tadej Droljc, who presented Singing Sand 2.0[4] in 8K 3D technology. Experiments were also made by using laser scanning and the acquisition and rendering of point clouds. In Poznan, the cathedral was scanned and the resulting image was presented in various formats - from 8K, through 3D up to 360° at 16K resolution[5] [5]. The content was also created in cooperation with the BBC and ScanLab, which scanned the Great Pyramid in Giza[6]. On this occasion, an interactive application was presented, allowing for a virtual experience inside the pyramid.

Live experimentation with 8K media streaming over long distances

Immersify also touches the subject of high-resolution media streaming. In cooperation with the artists, two long-distance transmissions took place. In September 2019, on the occasion of the Ars Electronica Festival, the artists' performance was broadcast in 8K to Deep Space in Linz[7]. The accordionist and dancer with armbands, processing through dance the sounds coming from the accordion, were recorded with four 8K cameras connected with 3G-SDI links via BPU-8000 to the IP LIVE switching system and transmitted over a distance of about 1,000 km in real time. This advanced experimental setup enabled typical broadcast event production with four 8K 60fps live video sources. The video editor in Poznan had a possibility to switch streams and create views consisting of several streams from different cameras connected to the system. The resulting 8K video stream served from the master IP LIVE system was sent directly to the PC server equipped with Blackmagic DeckLink 8K Pro capture card. The audio part was a mix of traditional broadcast techniques and artistic effect tools. The accordion was microphoned by two DPA instrumental mics and connected to Midac PRO1 console. This converted the analog sound to digital Dante protocol over IP. It was then fed to a MAC computer with MAX real time audio manipulation software. That computer also received movement data over Bluetooth, produced by the dancers MYO bands. It allowed the software to manipulate accordion sounds directly impacted by the dance performance. The ready 5.1 audio signal was then transferred over Dante to three AES3 digital lines and embedded into SDI signal which was fed to the encoding Workstation.

In addition, another artist in Linz influenced with his gestures the visualization of sounds appearing in the background, on the 8K screen in Poznan. The interaction was possible due to a total end-to-end delay of about 1 to 1.5 second. For encoding of the 8K video, an own solution developed at PSNC was used. The uncompressed 8K video captured by Blackmagic DeckLink 8K Pro was encoded with HEVC codec using NVIDIA NVEnc framework and two NVIDIA Quadro P5000 cards. Encoded signal was streamed in real-time using RTP/SRT protocols to Ars Electronica Deep Space 8K in Linz. On site the stream was decoded using software HEVC decoder from Spin Digital and presented to the participants of a dedicated 8K streaming session in Deep Space 8K. The complete configuration was depicted on Figure 1

[3] Always dead and alive, Theresa Schubert, http://theresaschubert.com/artworks/art/always-dead-and-alive/

[4] Singing Sand 2.0, Tadej Droljc, https://immersify.eu/content-demos/singing-sand-2-0/

[5] From the Inside – A different view of the Cathedral in Poznan, https://immersify.eu/content-demos/from-the-inside-a-different-view-of-the-cathedral-in-poznan/

[6] The Great Pyramid in 3D, BBC Ancient Invisible Cities, https://immersify.eu/content-demos/ancient-invisible-cities-the-great-pyramid/

[7] 8K HEVC live streaming from Poznan to Linz, https://immersify.eu/news/8k-hevc-live-streaming-from-poznan-to-linz/

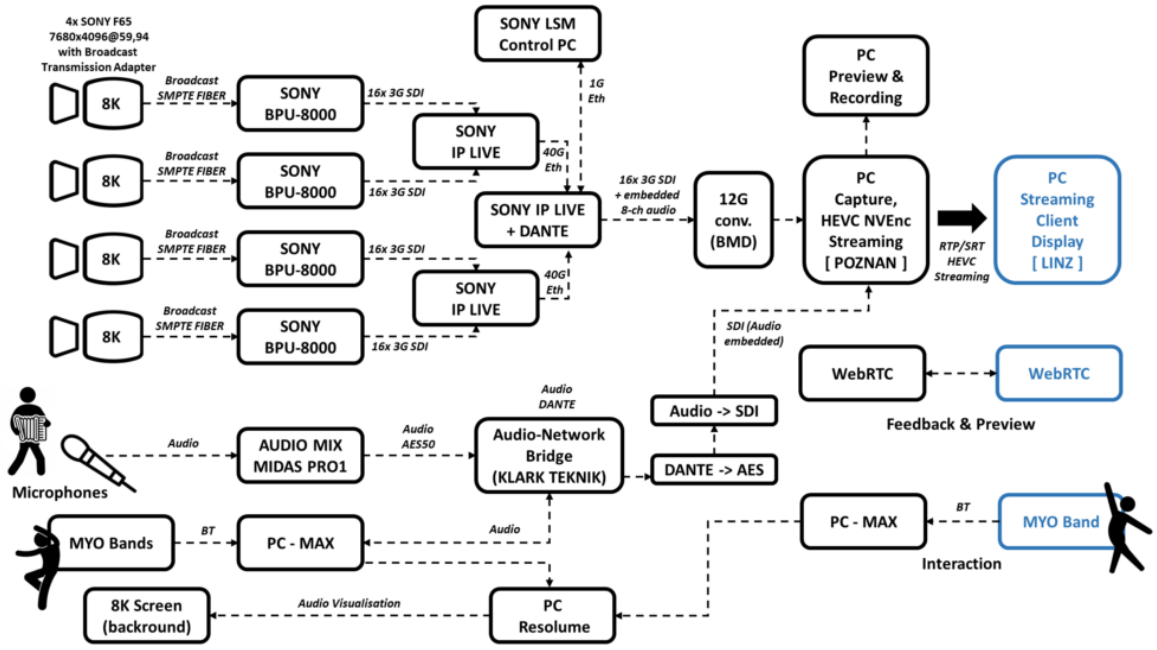

Fig.1. Hardware configuration of the 8K streaming from Poznan to Linz for Ars Electronica Festival

The second presentation took place on the occasion of InterBee exhibition in Japan[8]. In the PSNC studio in Poznan a young Polish artist, Anna Brzostek painting beautiful jellyfish was transmitted live to Tokyo to the NHK Technologies stand at the exhibition and displayed on 8K screen. In this case there were also four 8K cameras used, and one of them was installed on the robot arm, which allowed very dynamic, but precise movement.

For video switching, similarly to the previous case, the IP LIVE signal switching system was used, and for encoding we used the Advantech hardware encoder producing output HEVC stream at a bitrate of 80Mbit/s . As the Advantech encoder used a different video signal format as the IP LIVE system, for the compatibility between devices it was necessary to perform additional conversion of the output SDI signal (from SQ to 2SI). It was achieved with the PC equipped with 2 Decklink 8K Pro cards.

As a streaming solution the HLS protocol was used. This made it possible to use Amazon Web Services (AWS), MediaStore and CloudFront clouds for transmission, and streaming and displaying was carried out using Spin Digital software. Although the delay was much higher – 20 seconds, mainly because the cloud transmission for over 8,000 km, the 8K transmission could be viewed by multiple recipients simultaneously. At the main destination side, at the NHK Technologies booth at the InterBEE in Japan, a receiver using the Spin Digital HEVC media player (Spin Player) has been deployed. The player was connected to the 8K projector using a 12G SDI interface. What is worth mentioning here, the NHK Technologies with their partners used 5G connection for the last mile. The architecture of the demonstration is depicted on Figure 2.

[8] 8K Live Streaming from Poland to Japan at InterBEE 2019, https://immersify.eu/news/8k-live-streaming-from-poland-to-japan-at-interbee-2019/

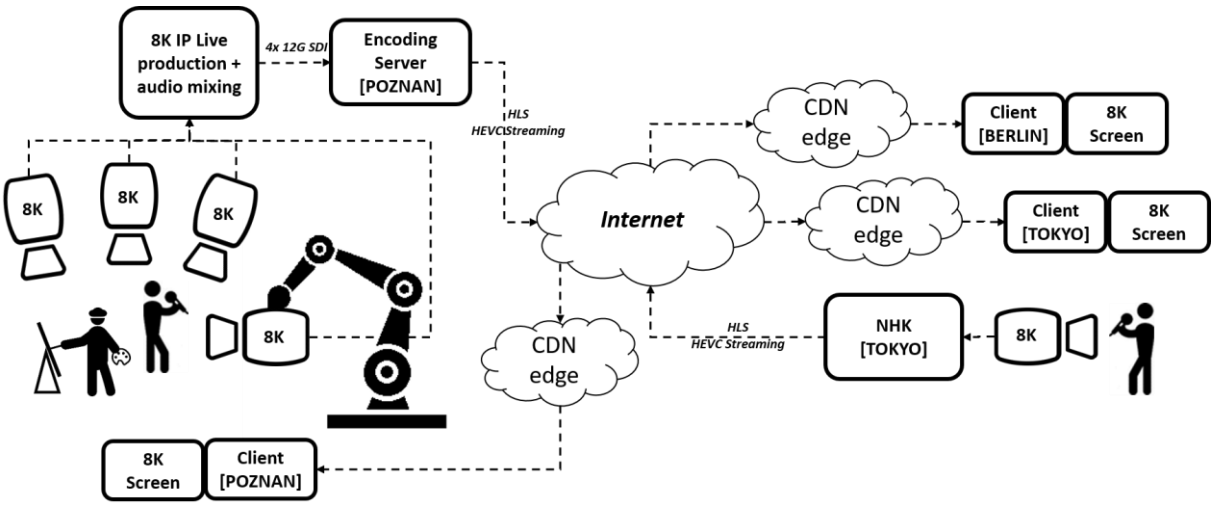

Fig.2. Hardware configuration of the 8K streaming from Poznan to Tokio for InterBEE 2019

References

[1] M. Alvarez-Mesa, S. Sanz-Rodríguez, C. C. Chi, M. Glowiak, R. Haring, Advanced video compression and rendering for highly immersive 8K+ applications, IBC19, Amsterdam 2019
[2] Immersify Project Web Site, www.immersify.eu
[3] M. Alvarez-Mesa, C. C. Chi, HEVC Software Media Player for Ultra-High-Quality Video: 8K, IDW18, Nagoya 2018
[4] K.Kurowski, M.Strozyk, M.Glowiak, M.Ciznicki, M.Alvarez-Mesa, B.Ludwiczak, A.Binczewski, PSNC advanced multimedia and visualization infrastructures, services and applications and beyond, WebVR 2018, Poznan, 2018
[5] M. Jaśkiewicz, J. Walczak, M. Glowiak, Immersify Guideline: Laser Scanning, https://immersify.eu/home/guidelines-reports/laser-scanning/

Authors' biographies

Maciej Glowiak has been working for PSNC since 2003, currently at the position of Head of New Media Department at PSNC. He leads research in areas such as 8K, 3D, 360° video, VR as well as uncompressed and compressed network streaming or immersive ambisonics sound. He is the administrative coordinator of the Immersify project.

Maciej Strozyk has been working for PSNC since 2003 as system designer and analyst for New Media Department at PSNC. His main research interests focus on the ultra high resolution (4K, 8K), multimedia streaming technologies and videoconferencing systems. Maciej has been participating as designer and developer in many R&D projects.

Szymon Malewski received the M.Sc. degree in Electronics and Telecommunication from the Poznan University of Technology in 2012. He has been working at PSNC since 2011 and he specializes in multimedia technologies, image processing and compression, as well as transmission systems and computer networks.

Dr. Mauricio Alvarez-Mesa is the CEO and co-founder of Spin Digital. Mauricio holds a degree in Electronics Engineering from the University of Antioquia in Colombia, and a PhD degree in Computing Engineering from Polytechnic University of Catalonia (UPC) in Spain. Mauricio has more than 15 years of experience in the field of video compression and high-performance computing. He has participated in several international research projects combining the media industry and the ICT sector. He is technical coordinator of the Immersify project.

Dr. Sergio Sanz-Rodriguez is a Senior Engineer and co-founder of Spin Digital. He received the degree in Technical Telecommunication Engineering in 2001, the M.S. degree in Telecommunication Engineering in 2005, both from Universidad Politécnica de Madrid, Spain, and Ph.D degree in Multimedia and Communications in 2011 from Universidad Carlos III de Madrid, Spain. He has more than 15 years of research experience in video coding and has published numerous articles in prestigious conferences and journals. He is also an expert in acoustics and audio processing, especially in the ambisonics field.