Long-distance Human-Robot Interaction with 3D UHDTV 60p video supported by VISIONAIR

Artur Binczewski[1], Maciej Glowiak[2], Bartlomiej Idzikowski[3], Maciej Strozyk[4], Eryk Skotarczak[5], Migel de Vos[6], Wladimir Mufty[7], Florian Draisma[8], Roy Damgrave[9], Marc Lyonnais[10]

[1,2,3,4,5]PSNC, Poznan, PL; [6,7,8]SURFnet, Utrecht, NL; [9]University of Twente, Enschede, NL; Ciena, CA

E-mail: [1]artur@man.poznan.pl, [2]mac@man.poznan.pl, [3]idzik@man.poznan.pl, [4]mackostr@man.poznan.pl, [5]eryk@man.poznan.pl, [6]migiel.devos@surfnet.nl, [7]wladimir.mufty@surfnet.nl, [8]florian.draisma@surfnet.nl, [9]R.G.J.Damgrave@ctw.utwente.nl, [10]mlyonnai@ciena.com

Abstract: **SURFnet, University of Twente, Poznan Supercomputing and Networking Center (PSNC) and Ciena performed a demonstration of human-robot interaction with stereoscopic UHDTV 60p video feedback over more than 2.000 km. The presented combination of a cutting edge technologies: all optical network, robotics, haptic force-feedback and ultra-high resolution compressed and uncompressed video, highlights the future potential of Human-Robot interaction in fields such as e-Health, e-Culture, e-Education and other real-time collaboration areas. UHDTV stereoscopic live video streams were transmitted through a 40G link established between Poznan, PL and Maastricht, NL with very low-latency. The demonstration took part during TERENA Networking Conference 2013 in Maastricht and was possible thanks to VISIONAIR project which calls for the creation of a European infrastructure for high level visualisation facilities that are open to research communities across Europe.**

Keywords: UHDTV, Human-Robot interaction, 40G network, High Frame Rate, Holography, Low-latency, VISIONAIR

1 INTRODUCTION

Solving a specialized or riskful mechanical problem from the other side of the ocean? Performing a specialized surgery over a large distance? Or just have social interaction by shaking hands remotely? Possibilities and opportunities are getting bigger, but are we ready? Common collaboration of three European research institutions: SURFnet, PSNC and University of Twente with Ciena as a network technology vendor proved that integration of different technologies in order to perform such experiment is possible. Partners devised a demonstration to show what is required to control a remote environment. There were several objectives of such a demonstration. The main goal was to reach out to researchers, teachers and students and create an interactive show using the newest network and visualization technologies, that would be easily understandable for all participants, even those not related to ICT. On the other hand, the experiment should use cutting edge visualization, haptic and network technologies. The first step was to find partners with specific competence in all required fields. This was possible through VISIONAIR project, which provides advanced visualization installations for researchers and users. VISIONAIR enabled the demonstration to be done during TERENA Networking Conference 2013.

The collaboration between NRENs, a university, several hardware vendors and the VISIONAIR project created the rare opportunity to showcase various innovations in one holistic demonstration. During the TERENA Networking Conference demonstration a lot of participants were able to play a game: ”the leaning tower of Pisa”. The participant locally controlled a robot arm to place small objects of different size and weight on the tower construction. This tower construction was located remotely in Poznan, Poland. The participants had several visual feedbacks such as: UHDTV 60p, holographic illusion and 2D and 3D displays through various compressed and uncompressed video streams. They were able to feel the haptic response from the robot arm.

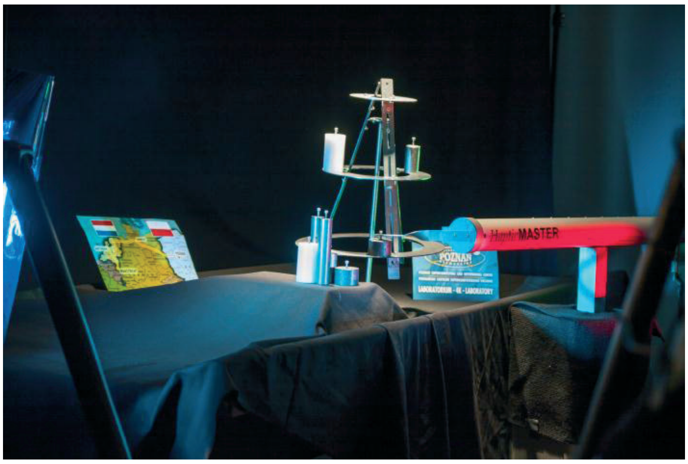

Figure 1: The leaning tower of Pisa game in Poznan studio driven by Haptic Master device

2 TECHNOLOGIES

To enable the demonstration, several innovative technologies were integrated and deployed together. The simplified demonstration scenario is depicted on Fig.3.

Corresponding author: Maciej Glowiak, PSNC, ul.Wichrowa 1a; 61-611 Poznan, Poland, +48 61 858 2024, mac@man.poznan.pl

2.1 Haptic Technology

To control a remote environment, a specific interface is required. A haptic robot arm gives humans the opportunity to make a virtual world tangible. Such an arm tries to mimic the real world. For this demo two Haptic Master devices from University of Twente were used. The haptic devices are able to perceive and transmit movements very precisely, and provide force-feedback. The person that controls one of the arms can feel weights, movements and pressure of the other arm. Both arms are network connected to each other and enable users to control precisely the position and movement of remote objects at a long distance with a very low delay.

2.2 Ultra High Definition video

Since the robot arms were not in the same room, even not in the same country, it was necessary to provide the users with visual feedback. By supplying different types of displays, resolutions and a low latency transmission of live video streams, users were able to watch the remote environment and interact with it. Participants could choose between large stereoscopic high frame rate UHDTV projection, 4K LCD displays as well as a holographic illusion display. The stereoscopic UHDTV projection at a rate of 60 frames per second was deployed using two professional 4K projectors (SONY SRX-T105) working together with polarization filters and a 5-meter-wide silver screen that gave an immersive experience to the audience. Another display was the holographic illusion booth. It was constructed using a 4K LCD display and a special holographic foil which was completely invisible to the naked eye and was placed in a special rig at 45° across the stage and then mirrored content off the LCD display. This view from the LCD was reflected upwards, reflected off the foil and gave the impression of a real 3D volumetric image on stage. As supportive devices three 4K LCD monitors from EyeVis and Astro Design were used as well as a 3D display from LG.

Video streams were transmitted from the PSNC studio in Poznan, where two 4K JVC (GY-HMQ10E) cameras on a special 3D rig were located and connected via special converters directly to Ciena's optical transmission system. Two separate live UHDTV video streams of 3840x2160 pixels enabled stereoscopic vision. As a support, several 2D and 3D camera streams were used in both directions to ease the experiment. All video signals from cameras were simultaneously streamed live without any video compression, however for quality comparison, one video signal was also encoded into JPEG2000. The total bit stream of transmitted video data reached about 30Gbit/s constantly over 5 days of the conference, what means that about 1,45 PB of data was transferred during the conference.

The latency added by the camera output-process, signal converting and displaying it through the projectors, was measured separately and took 3 frames (50ms). All devices in the chain: camera, video converters and projector were separated from the network and connected directly. Network transmission introduced additional 10 ms, so the total latency of 60 ms meant that the delay was not noticeable to any of the participants. They could experience the direct visual and haptic feedback that was in line with the movements and expectations of the participant controlling the haptic robot arm.

2.3 Uncompressed and compressed video

The demonstration focused on low-latency video transmission. In order to guarantee the maximal video processing speed the uncompressed streams were sent from cameras directly to the optical network system, from which it was retrieved and connected to the projectors. It caused the smallest delay between acquiring video and projection on the other side, but required perfect network connectivity and a bandwidth of about 24 Gb/s for stereoscopic high frame rate UHDTV. It's impressive, however very difficult to achieve in real applications. One of the purposes of the demonstration was to see how the usage of video compression could result in a lower bandwidth usage without losing the real-time and high quality experience. It was step forward to make human-robot interaction more applicable for regular usage in the field of research and education. For comparison, the JPEG2000 codec based on hardware IntoPIX's Pristine4 solution was used. JPEG2000 has been a Digital Cinema Initiatives standard for video compression since July 2005 commonly used in digital cinema and the movie industry for both – 2K and 4K images. However, now HEVC becomes a new standard for Ultra-High Definition video, the presence of hardware and software tools capable for real-time encoding and decoding, especially for 4K is very poor or none. For a couple of years PSNC and SURFNet have been successfully using the intoPIX system for 4K 3D live encoding and network streaming with many own implementations and improvements. That was the main reason of using JPEG2000 for comparison uncompressed and compressed video.

During the conference, two 4K displays, one next to each other, showed uncompressed and compressed video. JPEG2000 compressed video of a single 4K stream required 20 times less bandwidth and reached 500Mbit/s. There was no visual degradation of the quality, as JPEG2000 at such bitrates is perceived as a visually lossless codec. The only visible side effect of introducing compression to the chain was additional latency of 4-6 additional frames, what was caused by internal input/output queues of the encoding and decoding devices. Using a high frame rate of 60fps it means less than 100 ms which makes the fast and good quality video compression system still possible in real-time appliances.

2.4 Network for Video Transport

The live high quality uncompressed video feeds, in a combination with the haptic robot arm data and other supportive data (audio, videoconference) could generate a dataflow of 30 Gbit/s, therefore it was decided to configure a dedicated 40 Gbit/s all-optical path between Maastricht, NL and Poznan, PL.

Although the road distance between two locations is a bit more than 900 km, the optical link established via Poland, Germany (Frankfurt/Oder, Hamburg) and Netherlands reached about 2.000 km. The complete path was depicted on Fig.2.

An all optical dedicated path provided a low link latency connection (just 10 ms of one-way network latency!), required for almost real-time applications. Beforehand it was not sure whether a 40Gbit/s lightpath over a transmission distance of 2.000 kilometers was feasible. This was not only because of the huge distance, but also because SURFnet and PSNC are using different network vendor equipment. For experiment purposes Ciena provided an Ultra-Long Haul network system to establish the 40Gbit/s connection between SURFnet and PSNC. Error-free transmission of the 40G ULH alien wavelength using Ciena dual carrier BPSK coherent optical technology was deployed between Maastricht and Poznan. In total a distance of 2.000 km, mainly G.655 fiber, was covered without regeneration. Using SURFnet's existing Ciena uncompensated DWDM system between Maastricht and Hamburg and a compensated ADVA DWDM system between Hamburg and Poznan.

A major challenge with stereoscopic streams was that they consisted of several independent sub-signals (2 for HD 3D and 8 for UHDTV 3D). To keep these video signals in sync specific network hardware from Ciena was used. It was able to timestamp all Ethernet frames. This made it possible to guarantee perfect reconstructed, in sync, video streams. All the work with timestamping and synchronizing audio-video HD-SDI streams were done by Ciena network devices internally. The combination of the selected all optic network hardware and intensive communication between partner technicians gave this part of the project a unique and firm foundation that successfully enabled its extraordinary usage.

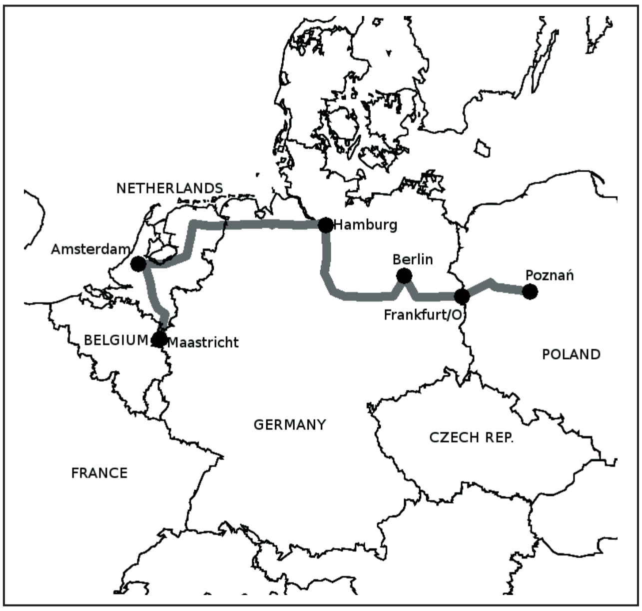

Figure 2: Optical path from Poznan to Maastricht

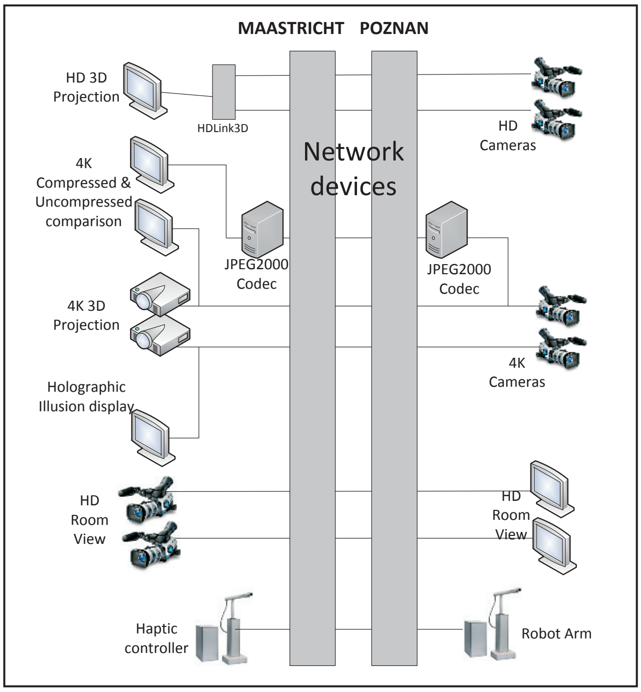

Figure 3: All technologies integrated in demonstration scenario

3 VISIONAIR SUPPORT

The preparation of the demonstration started several months before the conference. There was a lot of work to be done regarding the integration of video, network and haptic devices as well as establishing a dedicated optical path for the experiment. The physical access to PSNC and University of Twente visualization infrastructure was enabled by VISIONAIR project. VISIONAIR provides an open access to visualization infrastructure and services for the European research community. It comprises 4 main topics such as Scientific visualization, UHD networks, Virtual Reality and Collaborative Environments.

Scientists and researchers working on an interesting project related to visualization in one of mentioned areas are able to get access to one or several installations over Europe. Such access, even to the most advanced and innovative devices, is free of charge and a user may get back reimbursement for travelling and accommodation. It's also feasible for VISIONAIR partners to provide parts of their infrastructures at a distant location and then VISIONAIR project covers partial cost of transportation of scientific equipment. For human-robot interaction with UHDTV 60p video feedback the most relevant area was the UHD-NET activity, which focuses on tools and methods to visualize high quality images and allows sharing high definition images and movies over high bandwidth network. Visualization infrastructures from University of Twente and PSNC are part of UHD-NET activity, so SURFnet was able to apply for access and transportation of physical devices.

4 CONCLUSIONS

There were about 660 participants taking part in the TNC2013 conference. Most of them had an opportunity to play the game, successfully controlled the haptic robot arm. None of them experienced loss of video quality or video feedback delay as the total one way video latency took 60ms. Picture quality of all displays was excellent, however most of users particularly enjoyed an immersive stereoscopic large screen in UHDTV resolution. Users experienced the demonstration as if they were controlling a device in their local environment. Some of them even thought that the video screen was a high quality rendered video game, which shows the high quality and feasibility of the flawless and complete setup network, robotics, video and users that can actually use these together. The demonstration partners successfully combined multiple research and education aspects and will share this knowledge with others. Lessons learned from the experiment and the experience collected can be used in next collaboration scenarios for more complex use cases for e-Health, e-Culture or e-Education. The last but not least conclusion from the demonstration is that without VISIONAIR support such complex demonstration and collaboration of multiple partners would be much more complicated.

References

[1] UHDTV standard, ITU-R Recommendation BT.2020
[2] VISIONAIR project, http://www.infra-visionair.eu
[3] MUSION Eyeliner, http://www.musion.co.uk/#!/products-services/eyeliner
[4] The HapticMaster, a new high-performance haptic interface, http://www.eurohaptics.vision.ee.ethz.ch/2002/vanderlinde.pdf
[5] Binczewski, Głowiak, Idzikowski, Ostapowicz, Stroiński, Stróżyk , Research on stereoscopic 4K, TNC 2011, Prague
[6] Binczewski, Głowiak, Idzikowski, Ostapowicz, Stróżyk, New generation high resolution media in PSNC (4K/3D), Cinegrid, 2011
[7] Digital Cinema Initiatives specifications, http://www.dcimovies.com/
[8] JPEG2000 video codec, http://www.jpeg.org/jpeg2000

Figure 4: Maastricht demo room. From the left side: holographic-illusion display, haptic device stand, 4K 3D screen (rear), comparison of uncompressed and compressed video (right side).

Figure 5: Maastricht demo room. Haptic device in front, comparison of uncompressed and compressed video in the back

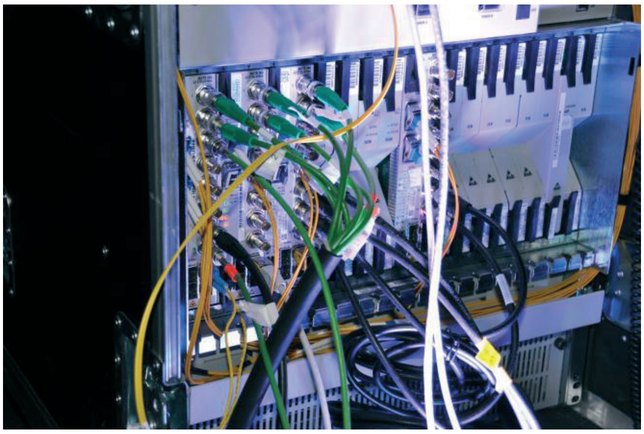

Figure 6: Ciena Ultra-Long-Haul for HD-SDI network transport

Figure 7: Poznan studio. Cameras and the Pisa game