


Live delivery of neurosurgical operating theater experience in virtual reality

Marja Salmimaa (SID Senior Member) 

Jyrki Kimmel (SID Senior Member)

Tero Jokela

Peter Eskolin

Toni Järvenpää (SID Member)

Petri Piippo

Kiti Müller

Jarno Satopää

Abstract — A system for assisting in microneurosurgical training and for delivering interactive mixed reality surgical experience live was developed and experimented in hospital premises. An interactive experience from the neurosurgical operating theater was presented together with associated medical content on virtual reality eyewear of remote users. Details of the stereoscopic 360-degree capture, surgery imaging equipment, signal delivery, and display systems are presented, and the presence experience and the visual quality questionnaire results are discussed. The users reported positive scores on the questionnaire on topics related to the user experience achieved in the trial.

Keywords — *Virtual reality, 360-degree camera, stereoscopic VR, neurosurgery.*

DOI # 10.1002/jsid.636

1 Introduction

Virtual reality (VR) imaging systems have been developed in the last few years with great professional and consumer interest.¹ These capture devices have either two or a few more camera modules, providing only a monoscopic view, or for instance eight or more cameras to image the surrounding environment in stereoscopic, three-dimensional (3-D) fashion. Most often, there is no provision for delivering the 3-D, 360-degree field of view in real time to the audience. The content is captured and the raw data is processed to present a uniform content scene to the users. This process involves “stitching” the views from the individual cameras together, which can be time-consuming. The stitched content is available for download, or for non-real-time streaming, after this processing step. The user is then dependent on various display media to view the content. These include computer monitors, where the full 360-degree field is visualized by mouse and keypad gestures; mobile phones, where the imagery is rendered by the user turning the device itself toward the desired direction of the view; or by dedicated eyewear that is able to render the 360-degree visual field in its entirety by the user wearing the eyewear turning his or her head in the desired direction. Especially the latter way of visualization has gained popularity in the last few years, and dedicated eyewear is available from many vendors. This eyewear either utilizes its own display or displays,² or it can utilize a mobile phone as the display device.³

In addition to pure entertainment, the eyewear and associated VR imagery can be utilized in professional use, such as training and education.⁴ Especially for those use

cases delivering a sufficient quality, VR experience is important yet challenging. In general, perceptual dimensions affecting the quality of experience can be categorized into primary dimensions such as picture quality, depth quality, and visual comfort, and into additional ones, which include, for example, naturalness and sense of presence.⁵

VR systems as such introduce various features potentially affecting the visual experience and the perceived quality of the content. First, the content can be purely synthetic, as in many gaming applications. Typically, such systems are able to deliver seamless 360-degree scenes, where, for example, the renderer determines the surface properties and amount of details in the 3-D models and structures thus influencing the naturalness of the content. Or the content may be captured for movie productions or created for documentary purposes, and captured by using ordinary or 360-degree camera systems. In those cases, the camera system specification determines the quality of the captured scenes, and especially in 360-degree camera systems, the stitching process may cause additional artifacts to the scenes. The most advanced systems may combine both synthetic and captured content or several different visual media objects as the system described in this paper.

Second, the visual display of the VR environment can consist of either 2-D (monoscopic) or 3-D (stereoscopic) images (or a combination). Usually, it is easiest to produce 2-D content with sufficient quality. With stereoscopic content, it is possible to create an illusion of depth in the content. Based on the earlier research, stereoscopic 3-D images can be perceived to some extent as sharper,⁶ and they potentially provide better image quality than

Received 02/17/18; accepted 03/01/18.

Marja Salmimaa, Jyrki Kimmel, Tero Jokela, Peter Eskolin, Toni Järvenpää and Petri Piippo are with the Nokia Technologies, Tampere, Finland; e-mail: marja.salmimaa@nokia.com.

Kiti Müller is with the Nokia Bell Labs, Espoo, Finland.

Jarno Satopää is with the Department of Neurosurgery, Helsinki University Hospital and University of Helsinki, Finland.

© Copyright 2018 Society for Information Display 1071-0922/18/2602-0636\$1.00.

conventional 2-D images.⁷ Also a number of studies have shown that compared with 2-D content stereoscopic, 3-D images have a greater psychological impact, that is, they enhance the viewers' sense of presence (provided that the depth magnitude is natural).⁸

Many of the perceptual dimensions affecting the quality of the experience are related to the sensation of depth in the content, and proper reproduction of depth can be considered as one of the most important aspects for the user experience. Still, most 3-D capturing systems are able to reproduce the scene depth properly only for certain depth ranges. If the depth range of the captured scene or the distance of the objects is not controlled, the objects located too close to the camera may be seen distorted in the captured content. In the worst case, the fusional mechanism of the visual system fails, and the viewer sees double images. Furthermore, introducing disparity as such does not automatically lead to a superior user experience. Stereoscopic or other image artifacts may have some implications to how users perceive or rate the quality. For instance, binocular image asymmetries may reduce the visual comfort^{7,9} and depending on the content (composition of the images), viewers may actually prefer non-stereo versions.

Third, the viewer can observe only a limited field of view at once, and when the viewer rotates his or her head, the system responds to that by rendering the content accordingly. For the feeling of being present in the environment rendered by using a VR system, the system needs to respond fast enough to the movements of the user. This requires low motion-to-photon latency which is sometimes hard to reach. High motion-to-photon latency may also cause some motion sickness and nausea symptoms for the user, thus making the viewing experience uncomfortable. Other factors contributing to the feeling of being present in the rendered scene include, for example, interactivity and the control over the task, multimodality in the presented content, ability to modify objects in the VR environment, the meaningfulness of the experience, and the realism and the consistency of the information.¹⁰⁻¹²

Finally, with streamed VR content, inefficiencies or limitations in the streaming systems may have some influence on the viewing experience.

Virtual reality (VR) technologies are emerging in the medical field in many areas, such as in VR therapy, planning of surgical procedures, and in many training applications.¹³ One of the emerging use cases for VR eyewear is, indeed, training in the medical field, such as in surgery. Education is expected to rise to the second-largest market position in medical and healthcare VR, only after therapeutic applications.¹³

This paper describes a VR imaging system and its use in live delivery of the stereoscopic 360-degree scenery from an operating theater to multiple users wearing VR eyewear. More precisely, the displayed content was a combination of the captured live 3-D 360-degree video combined with 2-D still images and live streamed 2-D video content, see Fig. 1.



FIGURE 1 — Partial view of an interactive neurosurgery experience from the operating theater.

The demonstration took place in June 2017 at the Helsinki University Hospital, where The 17th Helsinki Live Demonstration Course in Operative Microneurosurgery was concurrently organized.¹⁴

A Nokia OZO (Nokia, Finland) camera¹ with eight camera modules providing a stereoscopic 360-degree view was placed in an operating theater, the feed was transmitted to a server where the image stitching was performed in real time, and the imagery was delivered to a PC-based system with Oculus Rift eyewear, where it was possible to embed an interactive feed of the surgeon's microscope view into the VR view. As far as we are aware, in contrast to most stereoscopic 360-degree technologies, this was the first time an interactive live feed from the neurosurgeon's microscope camera was included in the VR view. In addition, for initial training purpose, a wireless local area network was provided to also transmit the stereoscopic 360-degree view as such to multiple viewers wearing Samsung Gear VR eyewear connected with a Samsung S7 smartphone. A user study was performed asking the participants fill in and return a questionnaire that contained a series of questions pertaining to the user experience, visual comfort, and display quality.

A system description of the trial is given, followed by a discussion on the user experience questionnaire. The results show that the participants were able to learn quickly how to use the player application. They also found the system practical and simple to use, and considered the virtual neurosurgery experience to be highly innovative and novel.

2 Stereoscopic 360-degree VR delivery system

The system for 360-degree VR delivery was built in the premises of the Helsinki University Hospital neurosurgery department, see Fig. 2.

2.1 Imaging subsystem

Nokia OZO camera was used for capturing a general stereoscopic 360-degree view of the operating theater. The main properties of the OZO camera are summarized

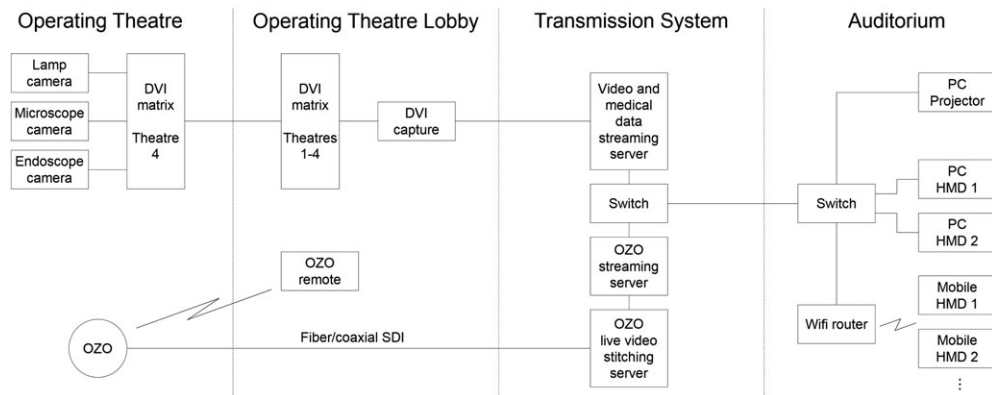


FIGURE 2 — Block diagram of the image delivery setup for the augmented VR demonstration.

in Table 1. Using dedicated hardware and software components for video processing, it was possible to deliver the imagery to the users from the OZO camera fully stitched in real time. The video quality of the live stream was mostly limited by the video stitching server. Offline optical flow based stitching can create very high quality stitching seams but is too slow to run with maximum quality settings for real-time applications. The stitched video was encoded in Ultra High Definition (UHD, 3840×2160) resolution using top-bottom stereoscopic format with equirectangular projection and real-time optimized stitching quality. Due to this standardized frame package size, the angular resolution in vertical direction is almost half (6 pix/deg) compared with the horizontal direction (10.7 pix/deg). The display resolution in the VR headsets would have allowed a use of higher resolution. However, in addition to the Oculus Rift based setup, also the Samsung

S7 mobile client could decode such video stream real time without problems.

In addition to the VR view generated by the OZO camera, other cameras integrated into the medical equipment in the operating theater were used for augmenting the VR scene. These included a camera providing the “lamp” view that renders a wide perspective image of the operating field, the “microscope” camera that captures the surgeon’s view through the operating microscope (the main view in microsurgical operations), and the “endoscope” camera used instead of the microscope for some operations. The operating theater personnel could select which of these cameras were streaming to our system as well as to the multiple display screens in the operating theater. The microscope camera was a Zeiss Pantero OPMI Trio 620 HD ccu 600 providing a 2k view to the area under operation. This camera provided the content to the live feed, augmented into the scenery provided by the OZO camera, when the user test participants were viewing the augmented 360-degree live feed on the Oculus Rift eyewear.

TABLE 1 — Basic properties of the OZO camera.¹

Property	Value	Unit/note
Number of sensors	8	Progressive scan global shutter
Sensor resolution	2048 by 2048	Pixels
Lens FOV	195	Degrees
Lens aperture	f/2.4	–
Angular resolution	6	′(Minute of arc)
Speed	30	Frames per second
Diameter	16.9	cm

FOV, field of view.

2.2 Image transmission subsystem

The OZO camera was located in the operating theater in two different positions, next to the wall and closer to the patient, see Fig. 3. Its image was sent as a 3G-SDI signal over an optical cable to a remote live video stitching server (acronyms presented in Table 2). The video stitching server



FIGURE 3 — OZO locations in the operating theater. On the left hand side the camera is next to the wall and in the middle figure closer to the patient. The ground plan of the operating theater on the right.

TABLE 2 — Acronyms related to the image transmission subsystem.

Abbreviation	Explanation
3G-SDI	Serial digital interface used for transmitting video signal
HTTP	Hypertext Transfer Protocol
MPEG-DASH	Streaming protocol that allows use of HTTP web servers and proxies
NGINX-RTMP	Open source component to enable conversion from RTMP to MPEG-DASH
PC	Computer with Intel x86/x64 processor
RTMP	TCP based Real-Time Messaging Protocol used for streaming low latency video
TCP/IP	Commonly used communications protocol used, e.g., on the Internet

combined the separate video streams captured by the OZO camera sensors in real time to produce a single 360-degree panorama stereo video stream. The stitching server was configured to send the resulting video as a RTMP stream over a local TCP/IP network to the streaming server. A NGINX server with the NGINX-RTMP module was used as the streaming server. The server converted the RTMP stream to the MPEG-DASH format used by the OZO SDK player and forwarded the stream to all player devices.

Video streams from the lamp, microscope, and endoscope cameras were streamed using another NGINX server. This server also had the HTTP module installed. It was used for providing medical imaging data selected by the surgeon and other 2-D images (e.g., user instructions) to the player application. The same server also managed the configuration files for the player applications. All servers were accessible both from the operating theater and from the demonstration area that was more than 100 m away. All traffic was delivered over a dedicated TCP/IP network that was separated from the main hospital intranet and the public Internet due to security reasons. Two sets of PCs with Oculus Rift eyewear were used to run the player application and to show the live augmented VR view of the operating theater to remote viewers.

2.3 Player software and display subsystem

A dedicated player application was developed with support for different VR displays. The player was implemented using the Unity 3D game engine. This game engine was chosen because it supports the Nokia OZO player SDK that was used for rendering the VR background video. It also has a built-in extendable system allowing users to interact with augmented objects in the VR scene. Further, the game engine allows targeting several VR platforms with minor or no code changes.

We prepared a player with basic 360-degree viewing capability for the mobile Samsung Gear VR devices. This setup was used for making participants familiar with the concept of live-streamed VR video. The main player application with more features was however run on PC-based hardware with Oculus Rift VR eyewear and hand-held Touch controllers. On top of the VR video, a

selection of augmented visual elements were displayed, including images, image stacks, and video streams. Medical imaging data of the patients and other 2-D still images were drawn to simple 3-D image planes using standard billboard technique. Billboards with movie textures were used for embedding the lamp/microscope/endoscope camera streams to the VR scene. The player application read a configuration file from the server that defined the image files and video streams, as well as their initial spatial locations in the scene.

For interaction, scaling and positioning of image and video billboards was supported. The billboard size was scaled instead of moving the billboard plane in depth dimension because of the possible convergence issues that could make the viewing experience uncomfortable. The user could select any of the floating billboard objects with the controller and freely change its size and position. For image stacks, browsing the images with left/right keys of the controller was also supported.

Figure 1 illustrates the VR player view on the Oculus Rift display. The observable field of view of the Rift depends on the user. The close to rectangular left and right display areas do cover diagonally an angular field of almost 110 degrees, binocular coverage being slightly less. However, due to the very limited exit pupil of the optics, as is common with most VR displays, edges of the field of view get blocked. Depending on the distance between the eyes and the optics, users can typically observe a round angular field of around 80–90 degrees. The floating image and video objects supported viewing of high-resolution content (e.g., 1080p), and their sizes were adjustable by the user to fill even the whole viewport, though initially their angular sizes were set to cover around 20–30 degrees.

3 Experimental setup and participants

The participants in the experiment were received in a small conference room at a different location in the hospital than the operating facilities. They were assisted to put on the eyewear (Fig. 4). They were first familiarized to the concept by letting them view the real-time operating theater view with the Samsung Gear VR with Samsung Galaxy S7 smartphone displays. The viewing device for the actual user study was the Oculus Rift eyewear, equipped with a handheld controller. The participants were instructed to turn around and explore the environment. They were also instructed to use the Oculus controller device that was operated by the user wearing the device in their right hand. With this pointer, it was possible to move and resize the augmented content, as the user wished. The participants were allowed to view the interactive real-time neurosurgery experience as long as they wished. After the participants had completed viewing, they were asked to fill in a questionnaire regarding the experience. Altogether, 22 neurosurgeons (17 men and 5 women) completed the questionnaires fully and



FIGURE 4 — Participants using the system delivering a live interactive neurosurgery experience from the operating theater.

provided their opinions on the experience. The participants represented different nationalities from around the world.

The mean age of the participants was 39 years and they had on the average 9 years' experience in neurosurgery. All had normal vision or corrected to normal vision (eyeglasses or contact lenses). Most participants viewed the interactive neurosurgery experience for approximately 5 min, while some participants who were interested in the experience viewed it for significantly longer periods of time (from 10 to 60 min).

4 Results

Overall, the experiment produced positive results. The participants rapidly learned how to use the player application and found the system to be practical and simple to use. Many participants commented that the system allowed them to experience the operation almost as if they were present at the operating theater. In general, the participants considered the virtual neurosurgery experience to be highly innovative and novel. They also described the experience as attractive and professional. The participants saw many possible applications for the real-time VR streaming technology in both medical education and practice.

The participants evaluated the image quality, perceived depth, naturalness and overall viewing experience on a seven step scale from “bad” = -3 to “excellent” = +3, see Fig. 5 for the results.

In general, it seems that the participants evaluated the quality of experience positively, as all the measured perceptual dimensions were rated clearly positive. The mean opinion scores and the standard deviations for the Image Quality and the Perceived depth were 0.95 (1.13) and 1.55 (0.96), respectively, and for Naturalness and Overall Quality 1.50 (0.91) and 1.59 (1.14), respectively.

The results of the evaluations of the Colors, Contrast, Sharpness, and Distortions can be seen in Fig. 6. The mean

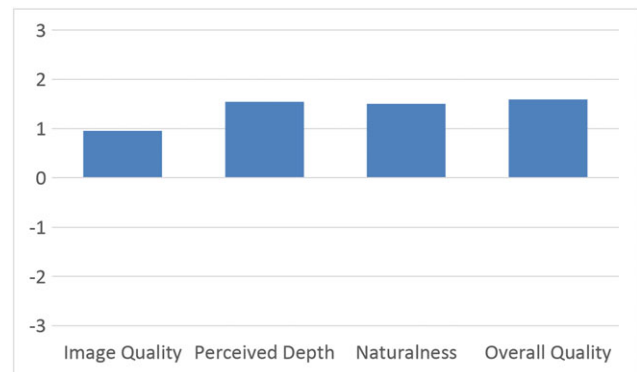


FIGURE 5 — Mean opinion scores for image quality, perceived depth, naturalness, and overall viewing experience.

opinion scores and the standard deviations for the Colors of the VR view and the Close-Up view were 1.50 (0.96) and 1.63 (0.95), respectively. The mean opinion scores and the standard deviations for the Contrast of the VR view and the Close-Up view were 1.40 (0.95) and 1.59 (0.91), respectively. The mean opinion scores and the standard deviations for the Sharpness of the VR view and the Close-Up view were 0.55 (1.30) and 1.27 (1.12), respectively. And finally for the Distortions the mean opinion scores and the standard deviations for the VR view and the Close-Up view were 0.72 (1.45) and 1.13 (1.32), respectively.

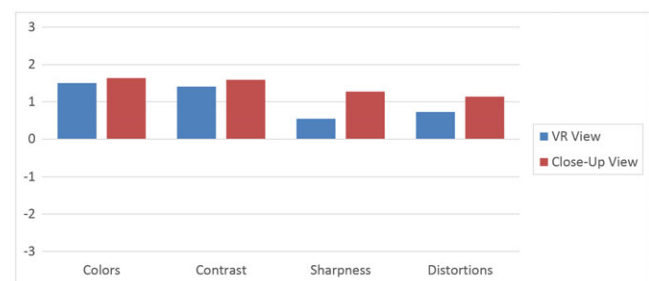


FIGURE 6 — Mean opinion scores for the colors, contrast, sharpness, and distortions evaluations for the VR view and the Close-Up view.

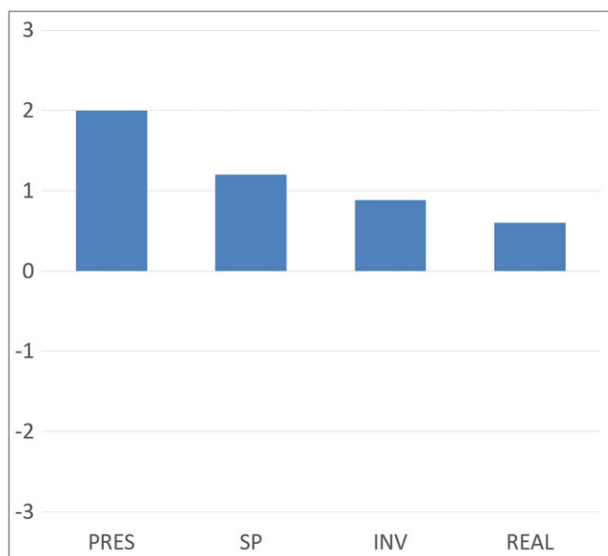


FIGURE 7 — Results of the igroup Presence Questionnaire. PRES, general presence; SP, spatial presence; INV, involvement; REAL, experienced realism.

Comparison of the Sharpness and the Distortion scores seems to indicate that in these dimensions the Close-Up view (augmented objects) was evaluated better. However, because of the nature of the data and the experimental setup, comparative statistical tests were not used in the analysis. We can still speculate what might have been the cause. The UHD frame size for the stereoscopic top-bottom format was clearly limiting the VR view (vertically down-scaled to half), while the Close-Up view was limited only by the VR display system. One consideration that would be of interest is how much the sacrifice in resolution to obtain stereoscopic presentation would affect the results compared with full UHD resolution 2D VR view.

In addition to the opinions on the visual experience, the sense of presence was measured by using igroup Presence questionnaire (IPQ).¹⁵ The results of the IPQ are reported in Fig. 7.

The mean opinion scores and the standard deviations were as follows: General presence 2 (0.81), Spatial presence 1.2 (1.07), Involvement 0.88 (1.31), and Experienced realism 0.60 (1.04). Based on the IPQ results, the participants clearly experienced a sense of being present in the operating theater.

5 Discussion

For the first time, participants in the Helsinki University Hospital Live course were able to interactively and remotely participate in the neurosurgery demonstrations. The difficulty so far has been that only 5–10 course participants can attend the surgery in the operating theater itself. The rest of the participants must view the operation on the projector in a lecture area organized in a lobby next to the operating theater. In principle, using the server and an

appropriate network connection, the neurosurgery operation could be broadcast worldwide using the technology presented in this trial. The participants evaluated the experience as positive, as the questionnaire summary shows, and the authors see value in developing the technology further. For instance, the augmented objects could include stereoscopic 3-D images and 3-D video, which is supported by advanced hospital imaging systems.

The system demonstrated in the study can be utilized in many fields of remote presence participation, medical surgery training only being a prime example of the capabilities. These fields include remote education and training in all medical fields for hospital personnel, medical students, and patients, field service, first-response assistance as well as on-call support.

Acknowledgments

The authors would like to acknowledge the support of Nokia Technologies' Digital Media Business unit, especially Olli Kilpeläinen and Peetu Hannukainen; Camilla Ekholm and Mikael Gustafsson from Nokia Technologies; as well as MD Ville Nurminen, MD PhD Miikka Korja, and MD Professor Mika Niemelä from Helsinki University Hospital and the administration of Helsinki University Hospital. The study was partly performed by funding from TEKES (Finnish Technology Fund).

References

- 1 J. Kimmel *et al.*, "Optics for virtual reality applications," Proceedings of EOSAM 2016, Berlin, 82–83 (2016).
- 2 Oculus Rift <https://www.oculus.com/rift/>
- 3 Samsung Gear VR <http://www.samsung.com/global/galaxy/gear-vr/>
- 4 Y. Pulijala *et al.*, "VR surgery: interactive virtual reality application for training Oral and maxillofacial surgeons using Oculus Rift and leap motion," in *Serious Games and Edutainment Applications*, M. Ma and A. Oikonomou, (eds.). Springer, Cham, (2017).
- 5 Subjective methods for the assessment of stereoscopic 3DTV systems, ITU-R Recommendation BT.2021 (2012).
- 6 M. Emoto and T. Mitsuhashi, "Perception of edge sharpness in three-dimensional images," *Proc. SPIE*, **2411**, 250 (1995).
- 7 F. Kooi and A. Toet, "Visual comfort of binocular and 3d displays," *Displays*, **25**, 99–108 (2004).
- 8 W. Ijsselstein *et al.*, "Perceived depth and the feeling of presence in 3DTV," *Displays*, **18**, No. 4, 207–214 (1998).
- 9 H. Self, "Optical tolerances for alignment and image differences for binocular helmet-mounted displays," Technical Report AAMRL-TR-86-019, Harry G. Armstrong Aerospace Medical Research Lab, Wright-Patterson AFB, USA (1986).
- 10 B. Witmer and M. Singer, "Measuring presence in virtual environment: a presence questionnaire," *Presence Teleop. Virt.*, **7**, No. 3, 225–240 (1998).
- 11 T. B. Sheridan, "Musings on telepresence and virtual presence," *Presence Teleop. Virt.*, **1**, No. 1, 120–126 (1992).
- 12 M. Lombard and T. Ditton, "At the heart of it all: the concept of presence," *J. Computer-Mediated Comm.*, **3**, No. 2, 20 (1997).
- 13 "Virtual reality in medicine and healthcare," Market report, ABI research (2017).
- 14 The 17th Helsinki Live Demonstration Course In Operative Microneurosurgery https://finland.aesculap-academy.com/go/?action=AkadEventData%26event_id=145706%26evdate=145713
- 15 T. Schubert *et al.*, "The experience of presence: factor analytic insights," *Presence Teleop. Virt.*, **10**, No. 3, 266–281 (2001).



Marja Salmimaa works as a Research Leader, Media Experiences at Nokia Technologies in Tampere. She holds an MSc (Tech.) degree from Tampere University of Technology. She has been a member of the Society for Information Display (SID) since 1997 and is currently a Senior Member and serving in the Program Committee of the SID Display Week. She is also a member of SPIE. Her research interests include new display technologies for mobile and wearable applications and related user experience, and augmented reality display systems as well as imaging technologies for virtual reality.



Toni Järvenpää received his MSc degree in computer science and engineering from Helsinki University of Technology in 2004. He is currently working as a researcher in the field of new emerging display and camera technologies and the related user experience in Nokia Technologies Finland.



Jyrki Kimmel is a Distinguished Researcher at Nokia Technologies in Tampere, Finland. He holds a DSc (Tech.) degree from Tampere University of Technology. He has been active in the Society for Information Display (SID) since 1999, currently serving in the Program Committee of SID and as Associate Editor of the Journal of SID. He is also a member of SPIE, IEEE, and EOS. His current research interests include new technologies for displays for mobile and wearable applications as well as imaging and display technologies for virtual reality.



Petri Piippo is a Principal Designer at Nokia Technologies in Tampere, Finland. He holds a Master of Arts in New Media degree from Aalto University. His current research interests include new technologies for virtual reality, especially in the area of user interface and interaction design.



Tero Jokela is a Principal Researcher at Nokia Technologies in Tampere, Finland. He holds an MSc (Eng.) degree from Tampere University of Technology. He is a member of ACM. His research interests include human-computer interaction and user interface software as well as multimedia applications.



Kiti Müller, MD, PhD, specialist in Neurology. Works as Senior Researcher in Medical and Neuroscience at Nokia Bell Labs. She is also Adjunct Professor in Neurology at Helsinki University Medical Faculty and Hospital.



Peter Eskolin holds a Master of Science degree in measurement technology. He has been working 17 years as a research engineer at Nokia Ltd. in Tampere, Finland. His expertise includes mobile communication technology, simulations, prototype implementation and user experience related applications and technologies.



Jarno Satopää, MD, PhD, works as a staff neurosurgeon at Helsinki University Hospital, Helsinki, Finland. His interests include bringing virtual reality technologies into professional use in the hospital environment.