



Telelife: The Future of Remote Living

Jason Orlosky^{1,2*}, Misha Sra³, Kenan Bektaş⁴, Huaishu Peng⁵, Jeeun Kim⁶,
Nataliya Kos'myna⁷, Tobias Höllerer³, Anthony Steed⁸, Kiyoshi Kiyokawa⁹ and Kaan Akşit⁸

¹Cybermedia Center, Osaka University, Osaka, Japan, ²School of Computer and Cyber Sciences, Augusta University, Augusta, GA, United States, ³Computer Science Department, University of California, Santa Barbara, Santa Barbara, CA, United States, ⁴Institute of Computer Science, University of St. Gallen, St. Gallen, Switzerland, ⁵Computer Science Department, University of Maryland, College Park, MD, United States, ⁶Computer Science Department, Texas A&M University, College Station, TX, United States, ⁷Computer Science Department, Massachusetts Institute of Technology, Cambridge, MA, United States, ⁸Computer Science Department, University College London, London, United Kingdom, ⁹Division of Information Science, Nara Institute of Science and Technology, Nara, Japan

In recent years, everyday activities such as work and socialization have steadily shifted to more remote and virtual settings. With the COVID-19 pandemic, the switch from physical to virtual has been accelerated, which has substantially affected almost all aspects of our lives, including business, education, commerce, healthcare, and personal life. This rapid and large-scale switch from in-person to remote interactions has exacerbated the fact that our current technologies lack functionality and are limited in their ability to recreate interpersonal interactions. To help address these limitations in the future, we introduce “Telelife,” a vision for the near and far future that depicts the potential means to improve remote living and better align it with how we interact, live and work in the physical world. Telelife encompasses novel synergies of technologies and concepts such as digital twins, virtual/physical rapid prototyping, and attention and context-aware user interfaces with innovative hardware that can support ultrarealistic graphics and haptic feedback, user state detection, and more. These ideas will guide the transformation of our daily lives and routines soon, targeting the year 2035. In addition, we identify opportunities across high-impact applications in domains related to this vision of Telelife. Along with a recent survey of relevant fields such as human-computer interaction, pervasive computing, and virtual reality, we provide a meta-synthesis in this paper that will guide future research on remote living.

Keywords: virtual reality, augmented reality, telelife, telepresence, human computer interaction

OPEN ACCESS

Edited by:

Frank Steinicke,
University of Hamburg, Germany

Reviewed by:

Hyowon Lee,
Dublin City University, Ireland
Richard Skarbez,
La Trobe University, Australia
Stefan Marks,
Auckland University of Technology,
New Zealand

*Correspondence:

Jason Orlosky
jorlosky-at-augusta.edu

Specialty section:

This article was submitted to
Virtual Reality and Human Behaviour,
a section of the journal
Frontiers in Virtual Reality

Received: 23 August 2021

Accepted: 25 October 2021

Published: 29 November 2021

Citation:

Orlosky J, Sra M, Bektaş K, Peng H,
Kim J, Kos'myna N, Höllerer T,
Steed A, Kiyokawa K and Akşit K
(2021) Telelife: The Future of
Remote Living.
Front. Virtual Real. 2:763340.
doi: 10.3389/frvir.2021.763340

1 INTRODUCTION

In 2020, the COVID-19 pandemic forced us to change the way we live. A major part of our daily lives became remote (Steed et al., 2020), and this change represented a major shift in how we work and socialize. For example, in many cases remote work can promote greater time flexibility, collaboration, and efficiency in working hours by reducing the need for transportation and allowing for multiple participants without the need of the same physical space. Consequently, the ideas of telepresence (Held and Durlach, 1992) and telework (Huws et al., 1990) regained momentum, presenting themselves as a meaningful way to overcome the limitations of remote work centered around video conferencing.

However, the promises of telepresence and remote interaction have yet to be achieved. The sudden, forced shift to remote operations without adequate preparation has produced an immediate need for new tools, research, and policies. Commercially available tools for remote operations inherently have many significant challenges, such as lack of 3D visuals and interactions. These

existing tools in the current transition towards remote technologies are in their infancy. With the recent changes, new issues related to ethics, social behavior, and accessibility have also arisen and are starting to be discussed among practitioners and researchers. For example, many users lost the close interpersonal feedback that they typically received through body language and embodied interaction with others. The loss of interpersonal feedback is partly due to the lack of high fidelity visual and non-visual feedback between themselves and their counterpart, sometimes resulting in a loss of attention and context-awareness critical for collaborative work and sometimes in an isolated or overwhelmed feeling in online settings.

This paper provides a of technologies encompassed in an ecosystem that we call Telelife. This represents an assembly of remote technologies and new innovative ways to enhance the way we live our intellectual, professional, community and personal lives in a remote context. This draws from prior work in telepresence, telework, and supporting technologies such as augmented reality (AR) (Azuma, 1997), virtual reality (VR) (Wang, 2002), pervasive computing, and relevant sub-branches. Telelife unifies these works into a single space, and we highlight recent state-of-the-art research that has brought society closer to remote living. While existing concepts often focus on the design of enabling technologies for remote collaboration or specific applications, Telelife offers a broader perspective that holistically addresses our needs as humans. As such, we revisit the need for new tools and concepts in remote operations, emphasizing the technologies necessary to support remote living in the next era.

Much like Sutherland's vision of the ultimate display (Sutherland, 1965), we present our idea of a possible future day, set for 15 years from now. We draw inspiration from existing product development cycles, similar to the process of monochrome phones in the pre-2000s turning into full-color

smartphones by the early 2010s, which is close to the 15-years timeline of our vision. While the ideas presented here may sound technologically utopian, our review of the state-of-the-art shows that much work towards this future is already underway and that it supports this vision.

Through several months of ideation, logical selection, and discussion between leaders in the field, we carefully formulated and refined the vision of Telelife. In addition, we identified domain experts from relevant fields and invited them to contribute. Lastly, we set out to support the vision by creating a meta-synthesis of the most relevant recent work. To do so, we gathered a database of all possible related references, focused on the past 3 years. We then logically selected those most closely related to this vision, but also included other significant core work from previous eras. Moreover, the vision and organization of this paper were further shaped by categorizing the related work we gathered.

Moreover, numerous issues related to user adoption of 3D technologies remain as major bottlenecks due to technical limitations, such as the lack of small form-factor near-eye displays (NEDs) or effective sensing devices. Other needs include the transmission and acquisition of realistic 3D graphics, which are significant challenges in computing and graphics domain. The greatest challenge for the realization of Telelife is to enable seamlessly integrated technologies in people's lives while providing experiences in line with human nature and psychology.

The key contributions of this work include the identification of challenges with realizing Telelife in various domains such as learning, privacy, and ethics, a synthesis of high impact research to help innovate the future of an increasingly remote society, and the identification of potential technologies and new applications that can be enabled as a result, as envisioned in **Figure 1**. Another goal of this synthesis is to help researchers find opportunities for growth in relevant fields and guide new research directions within the emerging Telelife ecosystem, for

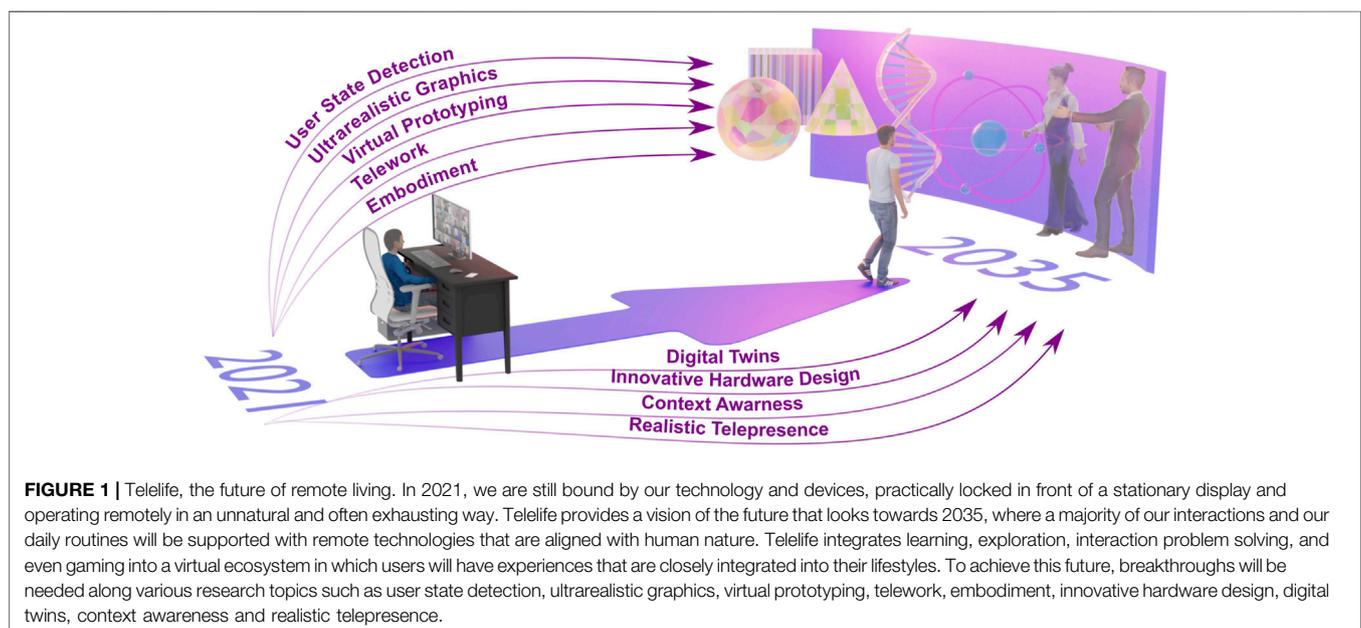


TABLE 1 | Telelife distinguishes itself from pervasive augmented reality, telepresence or virtual reality by providing a unique blend of remote technologies that offers new innovative ways to enhance the way we live our intellectual, professional, community and personal lives.

	Pervasive augmented reality	Telepresence	Virtual reality	Telelife
Use	Continuous	Sporadic	Sporadic	User aware
Control	Context controlled	User controlled	User controlled	Both user and context controlled
Interaction with physical environment	Visual (observing)	Bidirectional	Isolated	Bidirectional (physically)
Context of use	Context aware	Specific	Specific	Seamless
Ease of adoption	Learnable	Learnable	Application based	Inherent to human nature
Ease of use	Artificial	Inherent to human nature	Close to human nature	Inherent to human nature

which we also propose a set of grand technical challenges. More specifically, our contributions, which also represent the layout of the paper, are as follows:

- A unified research vision of Telelife. We help shape the vision for the near future by merging and unifying the highest-impact use cases and identifying interactions, technologies, and new paradigms within those areas. We name this agenda Telelife, a synthesis of discrete research visions.
- A day in 2035. We introduce a series of short speculations that exemplify a day in the Telelife of 2035 in various aspects of life.
- Challenges in Telelife. This section highlights existing challenges with remote interfaces, including smart homes, learning, collaboration, privacy, security, accessibility, adoption, and ethics.
- Innovative new research directions. Here we provide directions for researchers from a diverse set of fields to define the next steps in Telelife related research.
- Grand technical challenges. In this section, we group the challenges and research directions into succinct, high-level summaries to guide research from a holistic view.

2 THE VISION OF TELELIFE

For many years, researchers have dreamed of changing what we know as teleconferencing into telepresence. While this goal has merit, replicating a remote environment such as a classroom or office in its entirety still poses significant challenges. In contrast to the vision of telepresence, our synthesis consists of the technologies that will enable Telelife. The improvements outlined here will help remote computing environments to match the benefits of in-person meetings, and in some cases exceed these capabilities, as detailed throughout the paper.

Moreover, social media has had the unfortunate side-effect of reducing face-to-face meetings and interpersonal interaction. AR, VR, and related fields are currently in a unique position to “re-humanize” digital interactions between people. Though it has become possible for individuals to improve dialogue through messaging, images, and live video, we have yet to replicate the benefits of minute facial movements, eye contact, body language, emotion, and expression that make up a large portion of interpersonal communication. The goal is to bring back the human interactions not present in current technologies.

We distil Telelife down into four core areas: intellectual, professional, community, and personal life. Humans have a

natural tendency to explore, interact, and socialize in these contexts in 3D. What is essential is to focus on the right combination of research topics that will result in a state where the advantages of teleinteraction outweigh the disadvantages. It is not to say that 2D data representations such as spreadsheets or charts will become obsolete. However, Telelife will afford users a much richer experience and propel us past the cubicle, classroom, and desk paradigms pervasive in today’s world. Especially for small group interactions and creative work, 3D representations will provide the opportunity to replicate and exceed the experience of round-table discussions in physical space.

Additionally, just like collaboration platforms such as Slack or Discord have “channels,” Telelife will have new classes of 3D multitasking and messaging systems. Examples of this include leaving a 3D message or spatially visualizing channel data. These will become an essential part of the ecosystem that takes us into the next generation of 3D interfaces. More importantly, this ecosystem will need to ensure that new interactions are not invasive, forced, detrimental to productivity, or dehumanizing. Privacy, access control, and interaction should be managed with great care in order to ensure that new research adapts to the Telelife paradigm without hindering interpersonal interaction.

The ideas discussed in this paper build upon the previous work in fields of AR, VR, telepresence, communications research, and pervasive computing. In coming up with the vision of Telelife, we shed light on the most helpful and highest impact concepts from these fields to design the tele-ecosystem of the future. For example, while pervasive AR (Grubert et al., 2017) is embodied by continuous or always-on AR, Telelife is focused on the integration of information as a part of the user’s life experiences, socially, emotionally, and intellectually. **Table 1** provides a summary of the commonalities and differences between Telelife and other relevant fields.

In short, Telelife blends all aspects of the virtual spectrum Milgram et al. (1995) into the way we naturally live, focusing on the ecosystem in which users interact rather than specific applications. Previous visions such as virtual reality, pervasive augmented reality or telepresence do not offer a consistent vision towards re-humanizing remote interactions or making these interactions between people more natural. Our vision uniquely bridges the gap between real-life social interactions and technology. This paper also functions as a guide for researchers and industry to take into account and dedicate efforts toward development of 3D interactions, reconstruction techniques, and telepresence technologies that will propel society into a new age of remote living.

3 A DAY IN 2035

September 5th, 2035, Wednesday. Mike wakes up early, hoping to do something special for his son Jimmy's 10th birthday. While still in bed, he blinks to wake up his Telelife twin and asks him to prepare breakfast. Mike does not need to go to the kitchen personally. Instead, all available ingredients in the fridge are listed virtually in front of him with recommended recipes. His smart kitchen equipment automatically prepares the food upon his choice of ingredients, while Mike monitors the cooking progress through his twin's first-person view, and he chimes in to add a final touch of ground black pepper to his eggs.

It is now 8:30 am. Jimmy has finished breakfast and is ready for his school day. Every Wednesday is history day at teleschool, and Jimmy does not need to travel there physically. Instead, Jimmy switches the telechannel to "school" to meet with his teacher and friends in a fully immersive virtual space automatically customized for his room at home. As usual, kids in the classroom share 3D emojis that are shared only between their personal views so as not to interrupt others. In this virtual space, Jimmy is now sitting in an old-fashioned classroom for his history class with his classmates, each of whom designed their virtual uniform as homework for the visit back to early 21st-century. The teacher's avatar also takes the shape of an old-fashioned citizen to create the right ambience for the topic. "Today's class is about global pandemics. We are going to experience the Spanish flu of 1918 and the COVID pandemic of 2019–2021 first-hand. To better illustrate, I am first setting our digital environment time slider to the year 2020. Let us zoom into our current location to observe what was taking place in September." Jimmy begins his journey through a day in 2020 to feel the impact of the crisis and learn about the need for global preparedness.

While Jimmy is learning history at teleschool, Mike quickly browses today's news before setting off to work. The newspaper has long been outdated in 2035. Instead, the latest 3D reconstructed news scene is rendered as a small-scale model that travels along with Mike as he moves. He glances over today's stock prices, each showing historical balance sheet data rendered as a simple graph that an air-swipe of his finger can reorganize. At this moment, Mike is reading about a company specialized in personal garment manufacturing while feeling the fine texture via his virtual-haptic interface. While the 3D data shows cash flows and reduced debt, a recent fashion show sponsored by the company floats alongside. Within a short time, multi-channel and multi-model information is organized to help Mike decide whether he wants to buy a piece of the business.

Mike is now ready for work and decides to visit his physical office, though it has been quite a long time since his last visit. He has been working on a big architecture project in collaboration with a client in Asia. Today, Mike will demonstrate his model to his client using remote tele-fabrication. In his office, Mike first turns off the virtual Sun to sync up with the nighttime in Asia. He then logs into the digital meeting space with his clients. Mike and his clients are now co-located in the same virtual office in a split second, equipped with a tele-fabrication machine. Back in physical space, Mike's display can overlay authentic 3D visuals anywhere and at any size. His room is also fitted with a high-

fidelity, mid-air haptic feedback device synced to a display with less than a nanosecond of latency. In the physical space in Asia, his clients stand in front of a rapid prototyping machine that responds to Mike's remote actions in real-time. As Mike starts to explain the design concept while replaying his design process in the digital space, the 3D physical model is seamlessly built-in reality in front of his clients. His clients all keep their own small, physical copy of the model to take home and interact with later.

It has been a long day, and Mike's Telelife twin reminds him of Jimmy's birthday. Mike and his Telelife twin have been together since 2023. His twin has been his most immense emotional support, especially after Jimmy's mom passed away 3 years ago. After a long but productive day of telework across the globe, Mike leaves his office, sits in his self-driving car and browses Jimmy's 3D childhood photos and asks his Telelife twin for a suggestion on cake design. Twin-Mike suggests the one with Jimmy standing and smiling next to a giant furry monster from tele-playground. That was a somewhat lower resolution 3D-voxel map taken when Jimmy was three when voxel-based imaging was still in its infancy. Fortunately, Mike's digital twin has the latest AI retouching upgrades, so he produces a superresolution version for Mike to send to a local bakery for a customized birthday cake. Mike's twin's mainframe runs on the most recent photonic computer that runs on renewable energy resources. Though inexpensive twin-service is available, Mike maintains his twin's mainframe himself. It is set to auto-update with the help of a local chapter of the combined intelligence agency, in which he is also an active member.

It is now 6 pm. Mike is home and celebrating Jimmy's 10th birthday. Mike invited all his friends to Jimmy's tele-birthday and asked everyone to come up with their digital costume. They all talk, chat, draw and play around a large, interactive virtual table, and for many days following the event, Jimmy reminisces about the abounding laughter.

Fast forward to 2065. Jimmy has just turned 40, and he feels an internal warmth as he replays the moment from his 10th birthday in his mind using his direct brain-computer interface (BCI). To experience the old days, he decides to deactivate his BCI and enjoys his vacation on a sunny beach with his loved ones face-to-face for a time. "So much has changed," Jimmy says. So many discussions have taken place along the way to adopt these technologies, and so many political conflicts have been resolved by bridging the world through Telelife. In the end, Jimmy is glad that we have been able to re-humanize the way people live and communicate in today's teleworld.

4 CHALLENGES IN TELELIFE

Functional Telelife, as presented in the scenario above, faces several unique challenges, both technical and practical. While several of the technologies discussed are already available in some form, smooth integration between technologies, devices, and life presents many hurdles. Though much research has been done to address communication, interaction, display form-factors, participant visualization, and user feedback, much of the work done up to now has not yet become standard practice.

This section discusses some of the main challenges of achieving these in practice: interactions with intelligent homes, classroom and remote learning, collaboration with remote colleagues, and broader challenges related to privacy, security, ethics, accessibility, and adoption.

4.1 Smart Homes

Smart home technologies constitute digitally controllable systems such as lighting, heating and cooling, and Internet of Things (IoT) devices like kitchen appliances, robotic vacuum cleaners, and lawnmowers. For a smart home to function as envisioned, it must reliably interpret the inhabitants' movements and expectations, distinguish usual patterns from exceptions, correctly interpret inhabitant physiological signals and respond accordingly. The essence of a smart home is the interconnectedness of different automated systems (Norouzi et al., 2019). Immersive AR technologies can provide a unified cohesive UI layer for controlling these systems and services (Knierim et al., 2019).

However, fractured ecosystems, limited interoperability of heterogeneous technologies, and the lack of regulation and standards (Stojkoska and Trivodaliev, 2017) remain significant obstacles to the realization of this goal. In addition, the security and privacy concerns of the inhabitants may lower the acceptance of such technologies (Brush et al., 2011; Dahmen et al., 2017).

These challenges need to be addressed to support seamless integration of the smart home, IoT devices, and entities such as Mike's digital twin, as envisioned in the scenario above. Researchers should focus their efforts on the intersection of the immersive technologies that can control these systems with the protocols that interface the displays with other in-home devices. As mentioned later in the Privacy and Security section, security at each step of data flow, including networking, rendering, and application management, should be the focus of security research.

4.2 Learning

The introduction of rapid prototyping and immersive AR and VR technologies have enabled increasingly advanced forms of learning. High-quality interaction with the learning environment has been shown to improve learning (Dalgarno and Lee, 2010), and AR has been shown to outperform paper-based and digital methods of instruction delivery in terms of accuracy and time (Bhattacharya and Winer, 2019). However, these new learning styles are not yet widespread, partially because authoring content that is interactive, immersive, and educational is a complex process that lacks standardization and requires technical skills to create. While the use of AR-enhanced books is a powerful educational tool (Billinghurst, 2002), designing technologies for children, for example, requires taking into account their developmental abilities. However, there is relatively little work in that field. Moreover, great care should be taken to ensure that learning interfaces maintain current learning standards without introducing negative consequences. For example, many 3D interfaces that cater to visual learners may not be sufficient for auditory learners or those who prefer to read. This will require research into accessibility (part of the "Teleaccess" for all grand challenge), and the development of alternatives for different learning groups, such as virtual braille or supplementary auditory interfaces.

To realize the immersive remote learning scenario, we need to start with the senses. Sensory modalities including vision, audio, haptics, smell, taste, proprioception, and the vestibular sense present several actively explored challenges in ongoing research. For example, a display's field of view and resolution should match the human visual perception. We need improved spatialization and sound synthesis for audio and a wider variety of sensations for haptics. All these need to be integrated into a single system that enables the user to engage with digital content and remotely locate users as seamlessly as possible with physical objects and people in the same room.

Several recent survey papers review the state of the art in immersive learning opportunities, experiments, and benefits (Chernikova et al., 2020; Pellas et al., 2020; Hamilton et al., 2021; Kaviyaraj and Uma, 2021). Well-designed VR and AR educational experiences have been shown to raise students' engagement and learning. Settling on the most suitable formats and contents for specific educational goals is a significant context-specific challenge. With ever-increasing quantitative data from immersive learning case studies and deployed experiences, new opportunities arise for automatically recommending the most suitable formats for specific learning challenges.

4.3 Collaborating

Digital collaborative spaces can help realize mediated social experiences where distance disappears and users can interact as richly with remote participants as those in the same room. Collaborative spaces can be both 2D and 3D and take the form of video conferencing tools (e.g., Zoom, Gather Town), multiplayer video games (e.g., Fortnite), virtual office spaces (e.g., Arthur, Spatial) or virtual communities (e.g., VRChat, Microsoft Mesh). Collaborative virtual environments involve the use of networked systems to support group work (Dourish et al., 1996). The key concept behind these spaces is that each occupant is represented in a graphical form and can interact with the environment and the other occupants (Churchill and Snowdon, 1998). However, immersive virtual environments pose additional challenges to collaboration. For example, unlike video conferencing, increased bandwidth is required to transmit a 3D representation of a remote user's face and body. Though avatars can alleviate this problem, representations of minute facial movements in 3D are still a disadvantage that needs to be overcome.

An ideal collaborative space would track all user movement, include real-time facial capture for mapping speech, facial expressions and eye gaze for realistic interactions, and provide a space where users can move and engage with each other as if they were in the same room (Orts-Escolano et al., 2016). It is difficult because either users are burdened by significant amounts of equipment to enable tracking and dynamic scene reconstruction or the physical space needs to be instrumented, and both have their challenges. One considerable challenge in immersive collaborative environments is identifying the most suitable set of input techniques that are as efficient as the keyboard and mouse for desktop work. Effectively supporting the recording, tagging, transcribing, capturing, highlighting and sharing of content now popular with desktop videoconferencing tools and at the same time supporting the unique capabilities of

collaborative immersive tools (Sereno et al., 2020; Schäfer et al., 2021) are ongoing research challenges. Finally, the immersive nature of many co-located applications presents additional risks to the user, such as trip hazards, increased emotional responses to experiences such as acrophobia, and additional ways to take advantage of users through security vulnerabilities. The next section outlines some of these concerns.

4.4 Privacy and Security

As with other networked applications, privacy and security are paramount for remote interaction (Steed and Oliveira, 2009). With the increase in online interactions involving video or VR, we have seen incidents of Zoom bombing, fake identities, conference call hijacking, deepfakes, and other problems leading to unwanted disruption and distress. These issues often disrupt society both on a social and political level, which may be amplified in immersive experiences.

There are serious repercussions for privacy and security as some AR and VR companies start to connect how we move and interact in the virtual and physical worlds with how we think and feel on a personal level. While information about the physical environment is needed to provide the immersive experience or keep the user safe, aggregating that data with biometric and other data further exposes users and their protected data, such as medical records.

Because of the distributed nature and data flow in typical AR/VR systems, there is the danger that what is seen by the end user may not be intended by the developer since the rendered view may be vulnerable to alteration or injection by outside services or attackers. Each step of data flow, including detection, transformation, and rendering (i.e., inputs, operations, and outputs), should be protected, but each step requires different types of protections at the hardware, software, and communication level (Dahmen et al., 2017; Happa et al., 2019). In this respect, we need security and privacy standards for Telelife at a social and political level in the same way we have standards for wireless or network protocols.

Researchers have discovered vulnerabilities associated with some platforms (e.g., HTC Vive and the Oculus Rift) related to being able to change what the user is viewing, and also related to data flowing to and from the headsets (Yarramreddy et al., 2018). Changing what the user is viewing could lead to physical harm if safe movement boundaries are altered. By examining stored data on the VR platform, information about sites visited, time and date information, and user logs can all be recovered, similar to how file forensics of other system files can be recovered. Informing the user of safe or secure content in AR/VR is desirable, for example, by using a lock icon like that present in the address bar when visiting secure websites.

Physical characteristics of a user, such as head and hand position and orientation, can be used to identify users, leading to a complete effective loss of privacy, and the vast collected information about body language could even expose information and traits that the users themselves are at most dimly aware of (Bailenson, 2018). Miller et al. (2020) claim that 95% of users can be identified using this type of movement tracking information after only 5 min of gathered data for training the machine

learning identification system. Since movement information is normally captured for research purposes, standard data anonymization techniques will not prevent user identity from being determined from physical characteristics.

A recent literature survey (De Guzman et al., 2019) lists and reviews different protection approaches that have been proposed to ensure user and data security and privacy in mixed reality.

4.5 Accessibility and Adoption

According to the Center for disease Control (CDC), 1 in 4 US adults has some type of disability¹. With the influence of the world-wide COVID-19 pandemic on lifestyles, video conferencing has become a primary way to work, attend school, take care of fitness or socialize. Simultaneously, interest in video games has risen, offering an escape from the pandemic but also providing a new way to connect with family, friends and with colleagues². Video games are also increasingly being used for non-entertainment purposes such as education (Gee, 2003) and rehabilitation (Howcroft et al., 2012).

Unfortunately, special-needs populations, who might benefit significantly from remote and interactive technologies can often also be affected by hardware and software accessibility limitations. For example, one might imagine that the pandemic situation might have led to a certain amount of “leveling the playing field” for people with mobility limitations in the sense that now many more people were constrained to operate from their home environments. However, populations with limited mobility are also often underserved or overlooked as a target group, when it comes to the design of new immersive communication technologies such as VR (Mott et al., 2020).

Research has explored eye gaze (Liu et al., 2020), speech input, or mouth-based input devices with varying levels of success to enable users to interact with their computers, especially those with motor impairments due to amyotrophic lateral sclerosis (ALS), muscular dystrophy or cerebral palsy. While gaming systems³ and newer video games⁴ have come a long way in providing accessibility options, the limitations are mostly related to input devices. Newer technologies such as gaze-based methods or brain-computer interfaces (BCIs) have the potential to empower users. While there is ongoing research in these areas, much work needs to be done to make the devices as efficient and readily available as a keyboard and mouse.

A recent survey paper on accessibility research contributions at ACM CHI and ASSETS from 1994 to 2019 presents current and historical trends on the focus and potential biases of the research community trying to understand and support populations with disabilities and other access limitations (Mack et al., 2021).

4.6 Ethics

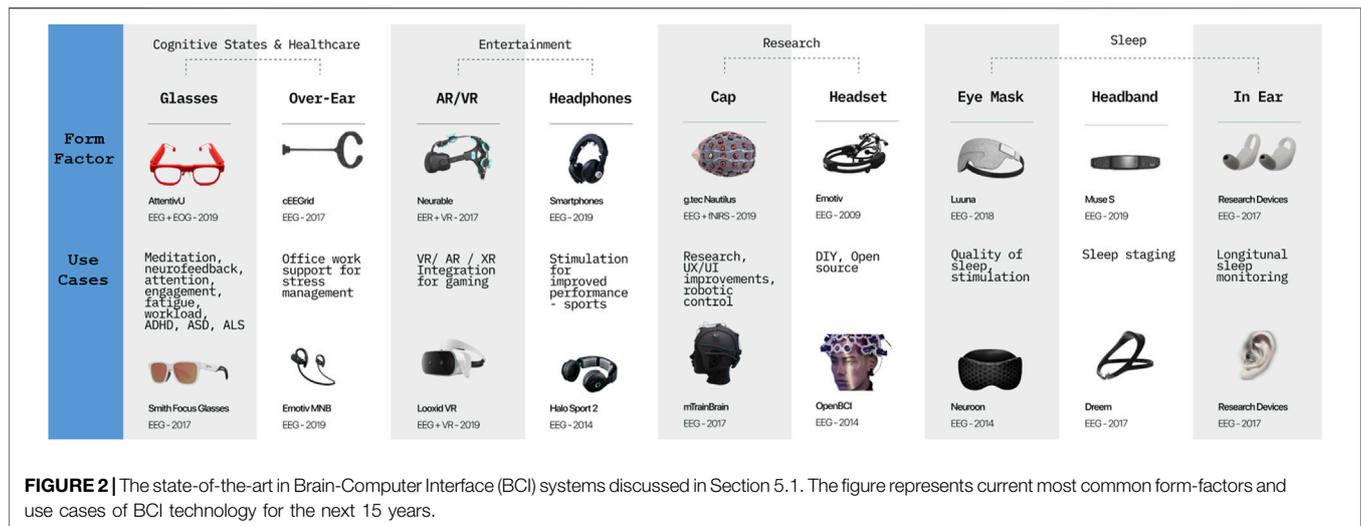
Historian Melvin (Kranzberg, 1986) first law of technology states, “technology is neither good nor bad; nor is it neutral.”

¹<https://www.cdc.gov/media/releases/2018/p0816-disability.html>

²<https://www.nytimes.com/2020/07/31/business/video-game-meetings.html>

³<https://www.xbox.com/en-US/community/for-everyone/accessibility>

⁴<https://steamcommunity.com/app/620980/discussions/1/1696046342864856156/>



The societal and personal implications of any technology ultimately depend on its use. For example, content moderation swiftly became a primary research topic in social computing, covering a variety of themes from toolkits to prevent email harassment (Mahar et al., 2018a), to studying unpleasant social behaviours such as cyberbullying in online games (Kwak et al., 2015). Social media platforms and remote collaboration tools that allow the use of text, image and 3D content sharing are also seeing increasing amounts of copyright violations, harassment and hate speech (Mahar et al., 2018b), which can be threatening on both a personal and cultural level. 3D experiences allow for more lifelike representations, so additional control should be afforded to users to moderate and regulate the content which they view or interact with. More recently, with the growing prevalence of AI technologies, the phrase “seeing is believing” is becoming less accurate, as videos, images and voice media are easily and effectively manipulated. The risks grow manifold as we move from the screen and begin to inhabit 3D spaces where an avatar may look and talk like someone you know, but is actually a 3D deepfake generated for malicious purposes.

Augmented Reality technologies add their own potential harmful consequences, including surveillance, peer surveillance and sousveillance, and even greater opportunities for bullying and other anti-social behavior (Heimo et al., 2014). Moderating content and regulating information policies is essential to prevent a greater potential for damage to social and cultural constructs in a fully connected telesociety.

5 INNOVATIVE NEW RESEARCH DIRECTIONS

The potential exists for a multitude of new experiences and interactions in our Telife vision. In this section, we lay out areas of research that are ripe for innovation and that will potentially have a high impact on the efficiency, engagement, and progress of remote living in the future.

5.1 Innovative Hardware Design

New hardware devices, some of which are shown in **Figure 2**, can pave the way to improve realism in our future experiences with Telife technologies while also providing a way to enhance software interfaces with a rich set of user data. In this section, we focus on innovative hardware designs for generating visuals and capturing a user’s attention and state. It should be noted that future-looking hardware designs go beyond the ones that we cover in this section as a perfect Telife experience would also require a multimodal hardware to support haptic feedback (Krogmeier et al., 2019), realistic audio (Zotter and Frank, 2019) and replication of olfactory experiences (Matsukura et al., 2013).

5.1.1 Hardware for Sensing User State

5.1.1.1 Brain-Computer Interfaces

There are different ways to measure user’s cognitive and mental states like attention, engagement or performance. A lot of studies and therapies turn towards more objective ways of measurement, in particular, physiological sensing. Typically, physiological sensing requires the user to wear a set of electrodes and devices that enable real-time monitoring of the user’s physiological and cognitive states. Examples include, but are not limited to the use of Electroencephalography (EEG)—non-invasive measurement of brain activity of the user; Electrooculography (EOG)—non-invasive measurement of eye movements of the user; Electrocardiogram (ECG)—non-invasive measurement of heart rate of the user or Electrodermal Activity (EDA)—non-invasive measurement of skin conductance. EEG and EOG can provide data relevant to information about attention and engagement of the user (Kosmyrna et al., 2019a). Electrocardiogram (ECG) and Electrodermal Activity (EDA) are often used to understand emotional arousal and to identify the magnitude of the emotional response (Hernandez et al., 2014). Electromyography (EMG) provides data on facial expressions linked to positive or negative valence (Sato et al., 2008).

Research in the field of BCI in particular, and physiological sensing in general, has been gaining momentum in the last 10 years, with the systems being used in rehabilitation (van

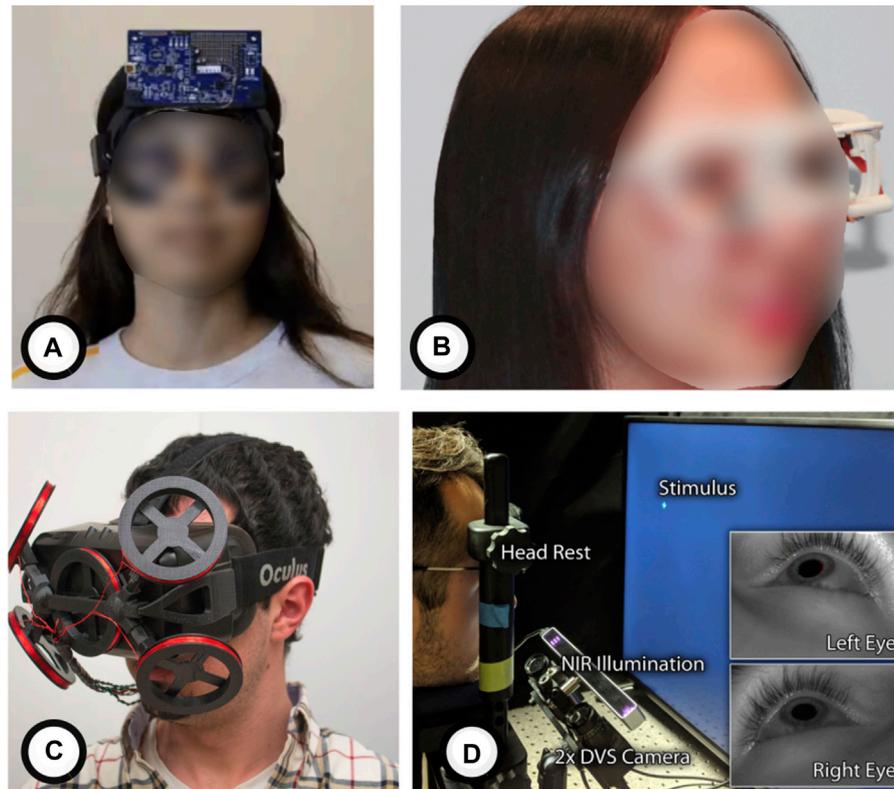


FIGURE 3 | State-of-the-art in the next generation eye and gaze tracking hardware discussed in Section 5.1: **(A)** Photodiode and LED based sensing (Li et al., 2020), **(B)** Camera based sensing (Lu et al., 2020), **(C)** Scleral-coil based sensing (Whitmire et al., 2016), and **(D)** Event camera based sensing (Angelopoulos et al., 2020). Though scleral-coil offers an invasive approach that is against our non-invasive Telife approach, it can potentially be a good match for contact-lens augmented reality near-eye displays at a possible future.

Dokkum et al., 2015), accessing mental states of the user (Afergan et al., 2014) and entertainment (Kosmyrna et al., 2015). However, as of today, these systems still remain expensive, bulky and uncomfortable as gel has to be applied to the electrodes, devices are wired and data is prone to classification errors due to the noisy nature of the signal. Thus, a lot of BCI systems nowadays are still being used in association with other input modalities like gaze trackers (Kosmyrna et al., 2015), or VR and AR near-eye displays (NEDs) (Amores et al., 2018). We believe that this trend will largely expand in the future, and in the next 10 years we will witness more NEDs with integrated hardware to measure brain activity. Beyond NEDs, neuromonitoring and neuroenhancing of one's wellbeing is expected to expand in the workplace for managing stress, meditating and measuring attention.

5.1.1.2 Eye Tracking Devices

Much like BCI systems, gaze interaction can potentially provide rich information for the next generation user interfaces. The most commonly used eye tracking hardware relies on video oculography, capturing videos of user's eyes using a set of cameras. Other examples of recent technologies are shown in **Figure 3**. While some of these prototypes are somewhat bulky and invasive (e.g., Scleral tracker), they will likely influence the

design of other unobtrusive eye trackers that might be integrated into thin contact-lens based displays (Chen et al., 2019) or yet-to-come technology. Although the resolution and the sampling rate of a camera are critical for the hardware to offer superior accuracy in eye and gaze location estimations, recent cameras that have high sampling rates and resolutions (Angelopoulos et al., 2020) can be demanding in terms of power and computational capabilities. These constraints may pose challenges in integrating such hardware with a wearable system such as NEDs (Stengel et al., 2015).

As an alternative to conventional cameras, the idea of *event cameras* has been demonstrated as a promising gaze tracking pipeline by capturing only the changing pixels in between frames (Angelopoulos et al., 2020). As it can get rid of needs in high-demanding computing power by capturing entire frames every moment (lesser pixels to capture), it could be as fast as 10 kHz. An interesting direction in eye and gaze tracking relies on using single pixel detectors (Topal et al., 2013; Li et al., 2020, 2017) that may potentially lead to all day useable eye and gaze trackers with low power and low computational complexity.

5.1.2 Display Technologies

In order to facilitate the next generation of Telife technologies, platforms that make use of NEDs (Koulieris et al., 2019) or 3D

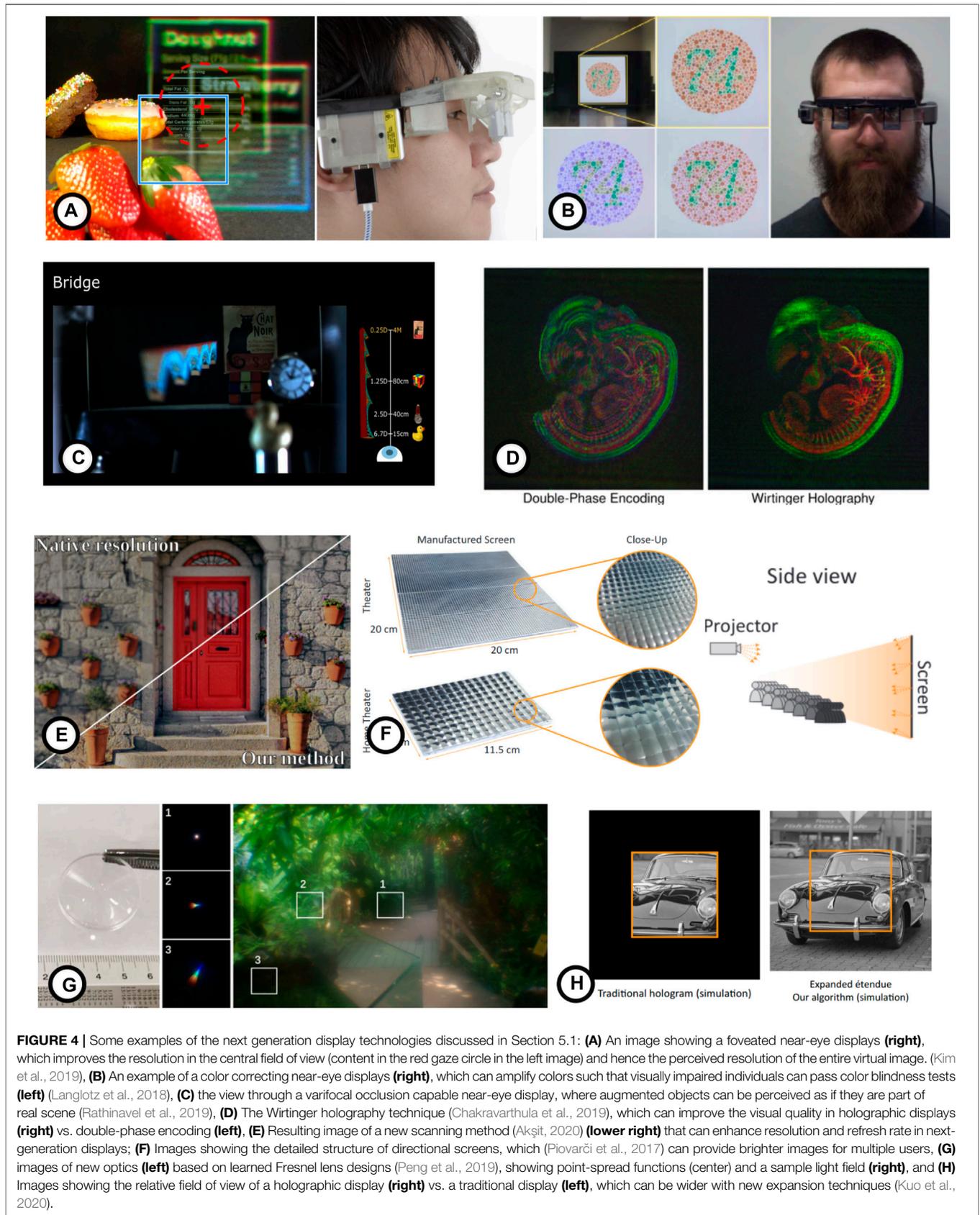


FIGURE 4 | Some examples of the next generation display technologies discussed in Section 5.1: **(A)** An image showing a foveated near-eye displays **(right)**, which improves the resolution in the central field of view (content in the red gaze circle in the left image) and hence the perceived resolution of the entire virtual image. (Kim et al., 2019), **(B)** An example of a color correcting near-eye displays **(right)**, which can amplify colors such that visually impaired individuals can pass color blindness tests **(left)** (Langlotz et al., 2018), **(C)** the view through a varifocal occlusion capable near-eye display, where augmented objects can be perceived as if they are part of real scene (Rathinavel et al., 2019), **(D)** The Wirtinger holography technique (Chakravarthula et al., 2019), which can improve the visual quality in holographic displays **(right)** vs. double-phase encoding **(left)**, **(E)** Resulting image of a new scanning method (Akşit, 2020) **(lower right)** that can enhance resolution and refresh rate in next-generation displays; **(F)** Images showing the detailed structure of directional screens, which (Piovarči et al., 2017) can provide brighter images for multiple users, **(G)** images of new optics **(left)** based on learned Fresnel lens designs (Peng et al., 2019), showing point-spread functions (center) and a sample light field **(right)**, and **(H)** Images showing the relative field of view of a holographic display **(right)** vs. a traditional display **(left)**, which can be wider with new expansion techniques (Kuo et al., 2020).

displays (Urey et al., 2011) will be of the utmost importance. However, form-factor, high resolution, low latency, and comfortable viewing experiences are still major challenges in adoption for such display hardware.

Like many other hardware design processes, designs of next generation display technologies can be enhanced with the use of machine learning techniques. In fact, some parts of this transition are already happening with new techniques to design diffractive optical components (Dun et al., 2020), fresnel lenses (Peng et al., 2019), and to drive holographic displays (Choi et al., 2020; Peng et al., 2020).

Either using machine learning or classical optimization methods, there isn't yet a single design pipeline that optimizes an entire design process of NEDs, while supporting custom needs such as providing prescription support (Chakravarthula et al., 2018), enhanced visual and color capabilities (Langlotz et al., 2018), and opto-mechanical designs that fit perfectly with a user's facial and eye structure. Beyond these, providing accommodation support in next generation displays regarding heating issues, occlusion support improved optics, dynamic range and field of view are still unresolved points in next generation display designs. There is also a noticeable recent work to tackle classical power and computational related problems in a near-eye display by actually removing active components from a near-eye displays (Itoh et al., 2021). We also provide examples of futuristic display technologies in the current literature as in **Figure 4**.

Fortunately, progress is being made in state of the art displays on form-factor (Kuo et al., 2020) and high resolution (Akşit, 2020). In the case of optical see-through displays, it is desirable to be able to change the transmittance of the screen, ideally on a pixel-by-pixel basis (Kiyokawa et al., 2003), and in the case of a video see-through display, it is desirable to be able to see the real environment with imperceptible motion-to-photon latency (Lincoln et al., 2016). Many research questions on the transitional interfaces, such as when and how to switch different spaces, and what kind of display hardware is optimal, remain to be answered.

5.2 Embodiment

5.2.1 Avatars and Agents

In addition to the technologies that enable Telelife, the ways that individuals are represented in virtual space are extremely important. Current practice with 2D teleconferencing is to stream 2D video and audio of conversation partners. The 3D equivalent is to stream a complete reconstruction of the individual with appropriate occlusion, resolution, lighting, and other visual characteristics (Orts-Escolano et al., 2016). However, many participants choose to represent themselves using an avatar, using an image filter, or even as a ghost-camera in the case of certain games. Though avatars are a well-studied field, other possibilities have been relatively unexplored.

For example, in a 3D online course with 1,000 students, it might not make sense to render 1,000 individual avatars in a single virtual room. Similarly, a teleconferencing window generally can't handle more than a few hundred participants. As a new direction in avatar development, avatars should take on more abstract characteristics. If a professor needs to view the number of students with questions, perhaps the students should be embodied as a question mark. In other words, rendering

someone according to their needs or current state may be more useful than rendering a facial expression.

Alternatively, it might be beneficial for a single simulated physical space to have multiple dimensions. Imagine a single, 20-person classroom that can hold 1,000 students in 50 different dimensions. The professor is rendered in all dimensions, but students pick one of 50 groups (dimensions) in which to be rendered. This would give participants the advantage of having an appropriately sized class in which small-group discussions could take place, but still allow for access of many more simultaneous attendees. This concept, which could be called dimension division multiple access (DDMA), similar to time division multiple access in networking, may be much more effective for scaling larger groups of people in experiences such as conferences, classes, or concerts.

In addition, intelligent agents will also play a key role in the hybridization and scaling of group experiences. Much like today's professors utilize teaching assistance to answer questions and grade work, tomorrow's tele-teaching agents could answer questions in real time during class without interruption, facilitate joint note taking between students, and reduce the teaching burden on educators. Similarly, an agent might assist with a presentation by switching slides, facilitating pointing, or waking sleepy participants. Individual feedback could also be customized and conveyed by a personal, trusted agent that empathizes with a user, thereby improving the chance that the user will adopt that feedback and improve his or her actions.

In addition, the geometry of the space in which the user operates in Telelife is generally different from that of the real environment around the user, except when the user is operating in the real environment as is without changing the position or scale, or when the user is operating in a remote or virtual environment with the exact same geometry as the surrounding real environment. Under such a situation, when the user moves around in the real environment, the geometry mismatch causes the user to collide with the walls or trip over obstacles. In order to solve this problem, there is some research that provides a virtual fence that is visible but impenetrable, or optimizes the relative positional relationship to maximize the range of movement. While there are many problems caused by the mismatch between real and virtual environments, there have been few studies addressing them. For example, there is limited research on the use of redirected walking to mitigate the mismatch between real and virtual environments (Sra et al., 2018) in telepresence. More studies are needed to alleviate the mismatch between real and virtual environments, as it is a key to the success of Telelife.

5.2.2 Multi-Modal Interactions

Multi-modal Interactions are the norm when interacting in the real world but the number of modalities usually decreases during digital interactions. For example, we communicate not only with our words but also with our bodies and facial expressions. While speaking, we can smell things in the environment, feel the movement of air, and sense the temperature among other

things. However, most of these sensory modalities are lost when communicating in a virtual environment, and recreating them is challenging. Aspects of conversation beyond speech, like facial expressions and body language are actively under exploration in addition to new forms of sensory feedback in immersive virtual experiences like temperature and wind (Ranasinghe et al., 2017), force feedback (Popescu et al., 2000; Choi et al., 2017), drag (Jain et al., 2016) and weight (Samad et al., 2019). There is some work exploring olfaction for PTSD therapy (Rizzo et al., 2010) and gustatory experiences (Narumi et al., 2011b,a) for attempting to include novel sensations in digital interactions.

Many challenges in hardware, sensory perception and software need to be overcome before these sensory modalities become commonplace in consumer virtual experiences. Researchers are starting to explore a new paradigm of human-computer integration that goes beyond human-computer interaction and explores how the muscles in the body and the digital interface form two types of closely coupled systems - symbiosis or fusion (Mueller et al., 2020). This new paradigm presents yet another unique set of challenges and opportunities in the form of technologies that go on, in or through the body, expanding interactions to newer dimensions. In addition to audiovisual information, multi-modal presentation is also a major technical challenge. The receptors for tactile sensation are distributed throughout the body, and the senses of smell and taste do not have primary components like the three primary colors of light (Miyashita, 2020). To share these senses with the remote environment, multi-modal sensing is required, which poses additional difficulties. There is some research on cross-modality, in which one sensory presentation is complemented by another sensory presentation, to generate different flavors through appearance changes (Nakano et al., 2019), but the effect is limited. Multi-modal sensing and display alone are extremely challenging already, but it will be even more challenging to realize multi-modal diminished reality. However, the ability to freely modulate these multi-modal information is crucial to enable flexible activities in remote and virtual environments with a high degree of reality, without making the user aware of the physical reality environment.

5.3 State Detection Using Biosignals

In order to develop convincing, valuable, and usable solutions, Telelife applications must be able to understand the user's needs, and by extension the user's state, including mental, physical, and emotional aspects. In this section, we will look into recent application level research for modalities such as eye tracking, BCIs and sensors that can help identify a user's state. Prior work has explored several paths to support users' cognitive and affective states. The four most common cognitive states the research community is focused on currently are engagement, attention, cognitive load and fatigue.

5.3.1 Eye Movements

Eye movements are an expressive method for detecting a user's intent, state, or context and they often precede physical action. Thus, gaze input can support (Fares et al., 2013) or substitute

(Templier et al., 2016) manual modes of interaction. According to Roy et al. (2017), eye fixations can be used to develop a model to predict the objects that a user is observing, which is a hint to how that user is interpreting a scene. Saccadic eye movements can enable predicting cognitive load, mental fatigue, attention, emotion, anxiety according to work by Duchowski et al. (2019). Using eye tracking, the work by Marwecki et al. (2019) introduces a technique for changing a scene dynamically in VR without being noticed by a user. In visual search tasks, similar methods, e.g., (Bektaş et al., 2019), (Gebhardt et al., 2019), can be useful for filtering redundant or irrelevant details.

5.3.2 Engagement

Engagement has mostly been studied and tested in learning environments: both subjective and objective methods have been used to provide information to teachers or presenters about the engagement levels of their students or audience (Hassib et al., 2017b; Kosmyna and Maes, 2019). In a recent study titled BrainAtWork (Hassib et al., 2017a), users were presented with their cognitive state, implicitly sensed using electroencephalography (EEG). The user's state was mapped to the workplace activities performed at the time. The authors used a visual modality as feedback about the cognitive state of the users, and the study was conducted in a lab. Users had the opportunity to see their engagement level both in real time as well as after the experiment was over. However, visual feedback might not be ideal for use in cognitively demanding environments such as workplaces, where frequent multitasking between digital windows or tabs within the same application may negatively affect cognitive performance. Research projects like *AttentivU* (Kosmyna et al., 2019b), a pair of glasses to inform the user about their engagement levels using haptic or auditory modality might be more suitable when designing future AR/VR applications for work/study environments in order not to overload the user's visual field, as well as considering real-time adaptation of the UI based on the physiological state of the user.

5.3.3 Attention

Methods for attention classification are currently not well adapted for workspace use and do not support discrimination between different attention types on a fine-grained level (Sohlberg and Mateer, 1987). Clinical Model of Attention discriminates between people's ability to maintain attention towards a single stimulus (sustained and focused attention); to switch attention between different stimuli (alternating attention); to pay attention to one stimulus while inhibiting others (selective attention), and to pay attention to multiple stimuli simultaneously (divided attention). This model highlights two challenges: quantifying attention (how much attention) and qualifying the nature of attention (what type of attention). Prior work on attention has shown that our well-being is tied strongly to our ability to manage attention successfully (Leroy, 2009). This creates an opportunity to design interactive systems that monitor and actively help users to manage their attention. The vision of pervasive attentive user interfaces encapsulates this well, stating that interfaces could "adapt the amount and type of information based on users'

current attention capacity, thereby simultaneously optimizing for information throughput and subtlety. Digital interfaces of the future could trade-off information importance with users' current interruptibility level and time the delivery of information appropriately" (Bulling, 2016).

5.3.4 Cognitive Load

Cognitive load has been measured traditionally either by standard questionnaires or by measuring a user's task performance. The NASA TLX is a common example of the first approach, where participants are asked to report their own cognitive load with regard to six different categories. Another example where study participants are asked to report their own estimates can be found in Sweller et al. (2011). One drawback of these approaches is that the answers are subjective. Furthermore, the self-reporting itself adds to the cognitive load. Measuring cognitive load through the performance in the task itself or in a secondary task (e.g., Lane Change Task for Automotive user interface, ISO 26022) only provides a rough estimate and is typically only suitable for laboratory studies and not for creating cognition-aware real-time systems. For interactive systems to be able to adapt their behavior accordingly, cognitive load information must be captured continuously and automatically—introspection is often not sufficient. Physiological sensors, such as EEG, and electrodermal activity (EDA) sensors show potential as possible solutions to this problem.

5.3.5 Fatigue

Fatigue can be defined as the unwillingness to continue performance of mental work in alert and motivated participants (Montgomery et al., 1995). It affects different cognitive functions including alertness, working memory, long-term memory recall, situational awareness, judgment, and executive control. Several technologies exist for monitoring fatigue levels, including eye tracking (Hutt et al., 2017) as well as video recording (Raca et al., 2015). These solutions are prone to errors and have limitations as cameras are sensitive to the ambient illumination and pose privacy problems, and are constrained to specific locations. Other measures include sensing physiological signals such as heart-rate variability (Boucsein et al., 2007), electrodermal activity (Byrne and Parasuraman, 1996), brain activity signals like EEG (Zander et al., 2010), Electromyography (EMG) (Fu et al., 2016), and Electrooculography (EOG) (Kosmyrna et al., 2019a). Current knowledge on the potential use of several types of sensors to measure each of these four phenomena in real-time on a fine-grained level is summarized in **Figure 2**. Despite multiple recent publications which investigate attention, cognitive load, fatigue and engagement, very few approaches have actually been deployed in the real world. The challenges to make this happen include the choice of modality or modalities with which to precisely measure one or more phenomena, hardware limitations (social acceptability of the device(s), form-factor, and comfort), and user value of these tools. Currently, passive sensing is only offered to the user, with very few options for adaptation of their environment and/or other active interventions.

5.4 New Applications

The research directions mentioned above will ultimately be driven by a set of applications within Telelife that call for the advance of technology. We present several of those applications below that will be at the forefront of remote interactions.

5.4.1 Gaze-Contingent and Context-Aware Assistance

Contextual information characterizes the interaction of a user with a computational system (Dey and Abowd, 2000). Context-aware systems provide information and services that are relevant to the current activity or task of the *user* and they are expected to adapt to dynamically changing situations of the inhabited *environment* and the *system* itself. According to Grubert et al. (2017), these three factors constitute the context sources. In a meeting context, audio and visual inputs to a computer can be disabled and tactile input (e.g., typing) can be enabled instead. In a training context, task-relevant visual content can be highlighted if it remains beyond the user's attention. An assessment of user and system behavior (preferably in real time) is a prerequisite to these adaptations. One way of understanding user behavior is to assess their visual perception through gaze-enabled systems. Eye tracking has recently become a promising feature in many HMDs because it enables foveated rendering, depth-of-field simulation, and studying the user's viewing behavior in the context of collaborative VR (Stengel and Magnor, 2016) and AR applications (Kim et al., 2019). By mimicking human vision, gaze-enabled systems can substantially improve rendering (Bektaş et al., 2015; Kim et al., 2019) and user performance (Bektaş et al., 2019). Such systems can direct user attention, augment their vision with contextually relevant information, and provide personalized assistance in different activities (Bektaş, 2020; Gardony et al., 2020). Some future implications of remotely performed activities may be beyond comprehension, so understanding how physical things and their digital counterparts inhabit our life will play a major role in future research (Mayer et al., 2018).

5.4.2 Digital Twins and Digital Companions

The future holds potential to be a hybrid one that seamlessly decorates our physical world with digital information (Weiser, 1999; Abowd, 2012; Ricci et al., 2015; Fuchsberger, 2019). In those environments the information from physical, social, and digital sources needs to be intertwined (Ricci et al., 2015). As we approach this hybrid future (**Figure 5**), we need to improve our understanding on how we can make use of various context sources, including humans themselves.

The digital twin of a physical system accounts for its real-world interactions between the environment and physical models through realistic simulations to update changes accordingly (Glaessgen and David, 2012). The twin is a living model that continuously adapts operational changes based on the collected data and simulation, either in real time or accumulated, and would allow forecasting the future of the corresponding physical counterpart (Liu et al., 2018). The



FIGURE 5 | Conceptual image highlighting a world in which the digital twins of people, buildings, vehicles, nature, etc. are integrated into the fabric of daily life.

twin will thus enable *mirroring* the life of its physical counterpart (Glaessgen and David, 2012). In our Telife vision, digital and physical beings (including humans) will grow with and adapt to each other. This adaptation needs to be carefully studied as it has the potential for expanding our perceptual capabilities while depriving others. To this end, monitoring behavioral changes will allow us to understand how users' cognitive abilities can be augmented in an environment with a digital twin. Remarkable applications of digital twins can be found in areas such as healthcare, manufacturing, and smart cities (Fuller et al., 2020). However essential questions such as how to generically model and build digital twins (Fuller et al., 2020) and how to standardize the communication protocols between physical and digital information resources need to be addressed (Tao et al., 2019).

Digital companions are informed by semantic representations of context sources (García et al., 2018), and can provide autonomous assistance to users by proactively acting, sensing, tracking, and communicating on their behalf. Recent work predicts that the decision support using digital twins will feed into these companions, which will be essential for industrial applications (Hartmann and Van der Auweraer, 2020) and possibly part of our daily life by 2030.

Despite some practical challenges, mobile and wearable sensors allow us to develop BCIs that are dedicated to monitoring a user's cognitive state based on psycho-physical measures. Such interfaces hold the promise to support operators, maintenance workers, and other users by eliciting device interactions in pervasive computing environments. In accordance with Shneiderman's proposal of having a human-centered approach in the design and development of next generation computational solutions (Shneiderman, 2020), the citizens of Telife may proactively remain at the center of the computation-loop, while advanced mixed reality (MR) interfaces support human-machine interactions and provide customized assistance.

5.4.3 Virtual Prototyping and Remote Fabrication

Telife will also benefit from advancements in other tools and systems that can physically produce and share virtual designs. An existing example is that of construction robots, which can build artifacts at remote locations where it would not be possible for a human supervisor to be safely present⁵. We have already seen a shift towards remote fabrication, with access to advanced personal fabrication tools (Dogan et al., 2021, 2020) (e.g., laser cutters, 3D printers) showing it is possible for open-source custom manufacturing in everyday life. Besides the growing number of consumer-grade fabrication tools and virtual resources, the functionality of software and hardware systems is also improving. Several advances (Peng et al., 2016; Kelly et al., 2019) support physical instantiation at much faster speeds. The progress of these technologies suggests a future where household hardware can be customized and created *in-situ*. This also highlights the need for education in design and fabrication skills, which will be a significant component of Telife.

Much recent research has made progress in enabling possibilities, such as RoMA (Peng et al., 2018), a system that combines 3D digital modeling with augmented reality and 3D printing. This supports seamless 3D construction by allowing a user to directly design in AR and print in real-time. In a similar context, compositional 3D printing (Kim et al., 2018) further promotes embodied fabrication using multi-modal input to support seamless construction of physical models during design. Scotty (Mueller et al., 2015), a self-contained appliance that allows relocating inanimate physical objects across distance, is another example of how remote physical editing will play a role in Telife. The user can send a physical copy to a remote place with the local one being destroyed, emulating the concept of teleportation, while at the same time demonstrating a creative way to protect copyrights. Scotty can be further equipped with AR

⁵<https://www.therobotreport.com/construction-robotics-changing-industry/>

editing features in which the teleported artifact can also be custom designed on-the-fly.

This body of work will be essential for remote collaboration. For example, during a digital meeting across the globe, one can design artifacts such as personal garments or large architecture at remote workshops. While the design could be fabricated locally by receiving real-time data, a consultant would have the opportunity to participate in the process, seamlessly updating custom requirements and other real-world constraints, visualized in the local space with a digital twin.

6 GRAND TECHNICAL CHALLENGES

Existing platforms for online ecosystems are in many ways well established, but they are still generally divided into separate fields. Here, we aggregate the most important challenges from these fields into unified grand challenges that represent the most important problems for the future of remote living. This is also designed to give researchers a holistic picture of the problem space for each area in order to better guide current and future research.

6.1 Re-humanized Teleinteractions

Tomorrow's technologies have the power to replicate the social and emotional interactions that we have lost in today's 2D world of social media. As such, we need to focus on re-enabling the face-to-face and interpersonal interactions that define human contact. The technology will not only have to be closely aligned with human nature, but will need to be accurate enough to reproduce facial expressions and body language. The grand challenge here is to reverse the process caused by 2D interfaces that has slowly detracted from interpersonal relationships and to re-humanize interaction so that remote living meets or exceeds the benefits of face-to-face contact.

6.2 Perfect Telepresence

An experience in which the physical and virtual are indiscernible from each other will require a fully integrated set of new hardware devices. These devices must render perceptually realistic experiences that support all senses (visual, auditory, tactile, smell, taste). To this end, true 3D displays will help unlock perfect visual experiences. Shape changing haptics will enable touch-experiences to render and augment physical objects of any shape or physical space. The combination with olfactory, auditory, and gustatory modalities will help reproduce experiences that can provide a true one-to-one correspondence to a user's physical body. The challenge with perfect telepresence is to address all of these sensory areas in a comprehensive display system.

6.3 Complete Cognitive Sensing

Understanding a user's needs, including emotional, social, intellectual, and innate, will enable interfaces that can better care for and facilitate communication in future societies. This refers not to the complete understanding of the human mind, rather the detection of the basic states such as stress, anxiety,

attention, fear, and others that can benefit from remote support systems. Two major research areas that show promise for the understanding of cognitive and contextual state include brain-computer interfaces and gaze tracking. Future components must be unobtrusive, support all-day usability, have seamless integration into eye- or head-wear, and require low computational and power resources. Artificial intelligence and machine learning will also be indispensable in this task, where the main challenge is to provide consistent, accurate estimates of the broad range of a user's cognitive states.

6.4 Technology Acceptance

Prior to the COVID-19 pandemic, some empirical evidence showed that the acceptance of wearable mixed reality technologies (e.g., headsets and smart glasses) was predictable based on the personality traits of the consumers (Rauschnabel et al., 2015). For instance, compared to the neurotic and sceptics the open and curious people were likely to adopt such devices. Furthermore, expectations of individuals from such devices (e.g., from being an innovative device to a fashion item) and social norms have an effect on the acceptance and adoption of such technologies (Rauschnabel and Ro, 2016; Herz and Rauschnabel, 2019). In accordance with our Telelife vision, the lock-downs caused by the pandemic might accelerate the adoption of emerging technologies and services which were previously perceived as nice-to-have, if not irrelevant.

6.5 Contextual Teleinterfaces

Our lives will be changed by future technologies, and our societies have to find an effective and beneficial way to cope with such changes per individual and for all members of society. From the design perspective, diverse virtual spaces and transitional interfaces (Grasset et al., 2006) that continuously provide a diversity in space types throughout a day need to be explored in more depth to switch smoothly from one task to another while maintaining cognitive continuity and physical safety, and not overwhelming users with monotonous experiences. Designs of these new experiences has to properly inform a user on what is virtual and what is part of real life, so that users can cope with the difference between virtual and real.

6.6 Teleaccess for all

A current major bottleneck in delivering online services is the lack of computing resources needed to provide such services to every member of society. Either in a local or cloud context, a wide variety of Telelife experiences must be supported with features that meet minimum standards of teleliving. Zero social isolation is a critical component in achieving this goal. The grand challenge here is to provide access, define minimum standards of teleliving, and design sustainable resources such that all individuals can live a full Telelife.

In summary, Telelife should be universally accessible, give us the best experience with the right set of technologies, understand our emotional and social needs, provide us with an interface that supports mental health and growth, and mitigate potential dangers to our wellbeing to the greatest extent possible.

7 CONCLUSION

The rapid evolution of technology has brought new opportunities for all of us. However, the pace of this evolution does not necessarily provide enough time to build a roadmap for the future ahead and seek means to pursue ethical designs and processes. Laying a high-level tangible roadmap can provide the opportunity to effectively solve problems by uniting research communities to move towards a common goal. We believe the current communities involved in virtual and augmented reality need such a roadmap to impact the way our societies live positively.

In order to ensure that remote technologies are well aligned with human needs, we introduce a vision called *Telelife*, a high-level term that encapsulates previously introduced visions and concepts such as telepresence, telework and AR/VR. We cherry-pick and merge the essential items from those fields, identify relevant challenges in achieving *Telelife*, provide an overview of the state-of-the-art, and outline future directions in research that will help re-humanize remote technologies and teleinteraction.

Remote living is a multi-dimensional problem larger than all of us as individuals and will require efforts from each of our unique research fields. We hope to unify the visions of researchers so that their work will better fit into the upcoming *Telelife* ecosystem. Moreover, the concepts presented here can inspire and guide future generations of research on remote living. Lastly, we aim to raise the opportunity to discuss future questions about how we can and *should* live our remote social, intellectual, professional, and personal lives.

REFERENCES

- Abowd, G. D. (2012). "What Next, Ubicomp? Celebrating an Intellectual Disappearing Act," in Proceedings of the 2012 ACM Conference on Ubiquitous Computing, New York, NY, USA (New York, NY: Association for Computing Machinery), 31–40.
- Afergan, D., Peck, E. M., Solovey, E. T., Jenkins, A., Hincks, S. W., Brown, E. T., et al. (2014). "Dynamic Difficulty Using Brain Metrics of Workload," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (New York, NY: Association for Computing Machinery), 3797–3806. doi:10.1145/2556288.2557230
- Akşit, K. (2020). Patch Scanning Displays: Spatiotemporal Enhancement for Displays. *Opt. Express* 28, 2107–2121. doi:10.1364/OE.380858
- Amores, J., Richer, R., Zhao, N., Maes, P., and Eskofier, B. M. (2018/2018). "Promoting Relaxation Using Virtual Reality, Olfactory Interfaces and Wearable Eeg," in IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN), 98–101. doi:10.1109/bsn.2018.8329668
- Angelopoulos, A. N., Martel, J. N., Kohli, A. P., Conradt, J., and Wetzstein, G. (2020). Event Based, Near Eye Gaze Tracking beyond 10,000 Hz. arXiv preprint arXiv:2004.03577
- Azuma, R. T. (1997). A Survey of Augmented Reality. *Presence: Teleoperators & Virtual Environments* 6, 355–385. doi:10.1162/pres.1997.6.4.355
- Bailenson, J. (2018). Protecting Nonverbal Data Tracked in Virtual Reality. *JAMA Pediatr.* 172, 905–906. doi:10.1001/jamapediatrics.2018.1909
- Bektaş, K., Çöltekin, A., Krüger, J., and Duchowski, A. T. (2015). *A Testbed Combining Visual Perception Models for Geographic Gaze Contingent Displays*.
- Bektaş, K., Çöltekin, A., Krüger, J., Duchowski, A. T., and Fabrikant, S. I. (2019). "Geogcd: Improved Visual Search via Gaze-Contingent Display," in Proceedings of the 11th ACM Symposium on Eye Tracking Research and

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This work was funded in part by the University College London and Osaka University partnership fund, grant #Na20990020, by the Office Of Naval Research Global, grant #N62909-18-1-2036, by the Basic Research Fund of the University St.Gallen, and by the Innosuisse—Swiss Innovation Agency (Award Number: 48342.1 IP-ICT).

ACKNOWLEDGMENTS

The authors would like to thank reviewers for their valuable feedback. Special thanks to Michael Nowatkowski from Augusta University for his input. Thanks to Nina Jay Cragg for helping create **Figure 2**.

- Applications, New York, NY, USA (New York, NY: Association for Computing Machinery), 1–10.
- Bektaş, K. (2020). "Toward a Pervasive Gaze-Contingent Assistance System: Attention and Context-Awareness in Augmented Reality," in ACM Symposium on Eye Tracking Research and Applications, 1–3.
- Bhattacharya, B., and Winer, E. H. (2019). Augmented Reality via Expert Demonstration Authoring (Areda). *Comput. Industry* 105, 61–79. doi:10.1016/j.compind.2018.04.021
- Billinghurst, M. (2002). Augmented Reality in Education. *New horizons Learn.* 12, 1–5.
- Boucsein, W., Haarmann, A., and Schaefer, F. (2007). "Combining Skin Conductance and Heart Rate Variability for Adaptive Automation during Simulated Ifr Flight," in *Engineering Psychology and Cognitive Ergonomics*. Editor D. Harris (Springer Berlin Heidelberg), 639–647. doi:10.1007/978-3-540-73331-7_70
- Brush, A. B., Lee, B., Mahajan, R., Agarwal, S., Saroiu, S., and Dixon, C. (2011). "Home Automation in the Wild: Challenges and Opportunities," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, NY, USA (New York, NY: Association for Computing Machinery), CHI '11, 2115–2124. doi:10.1145/1978942.1979249
- Bulling, A. (2016). Pervasive Attentive User Interfaces. *Computer* 49, 94–98. doi:10.1109/mc.2016.32
- Byrne, E. A., and Parasuraman, R. (1996). Psychophysiology and Adaptive Automation. *Biol. Psychol.* 42, 249–268. doi:10.1016/0301-0511(95)05161-9
- Chakravarthula, P., Dunn, D., Aksit, K., and Fuchs, H. (2018). Focusar: Auto-Focus Augmented Reality Eyeglasses for Both Real World and Virtual Imagery. *IEEE Trans. Vis. Comput. Graphics* 24, 2906–2916. doi:10.1109/tvcg.2018.2868532
- Chakravarthula, P., Peng, Y., Kollin, J., Fuchs, H., and Heide, F. (2019). Wirtinger Holography for Near-Eye Displays. *ACM Trans. Graph.* 38, 1–13. doi:10.1145/3355089.3356539

- Chen, J., Mi, L., Chen, C. P., Liu, H., Jiang, J., and Zhang, W. (2019). Design of Foveated Contact Lens Display for Augmented Reality. *Opt. Express* 27, 38204–38219. doi:10.1364/oe.381200
- Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., and Fischer, F. (2020). Simulation-based Learning in Higher Education: a Meta-Analysis. *Rev. Educ. Res.* 90, 499–541. doi:10.3102/0034654320933544
- Choi, I., Culbertson, H., Miller, M. R., Olwal, A., and Follmer, S. (2017). “Grability: A Wearable Haptic Interface for Simulating Weight and Grasping in Virtual Reality,” in Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology, 119–130.
- Choi, S., Kim, J., Peng, Y., and Wetzstein, G. (2020). Michelson Holography: Dual-Slm Holography with Camera-In-The-Loop Optimization. arXiv e-prints, arXiv:2009
- Churchill, E. F., and Snowden, D. (1998). Collaborative Virtual Environments: an Introductory Review of Issues and Systems. *virtual reality* 3, 3–15. doi:10.1007/bf01409793
- Dahmen, J., Cook, D. J., Wang, X., and Honglei, W. (2017). Smart Secure Homes: a Survey of Smart home Technologies that Sense, Assess, and Respond to Security Threats. *J. Reliable Intell. Environ.* 3, 83–98. doi:10.1007/s40860-017-0035-0
- Dalgarno, B., and Lee, M. J. W. (2010). What Are the Learning Affordances of 3-d Virtual Environments? *Br. J. Educ. Technology* 41, 10–32. doi:10.1111/j.1467-8535.2009.01038.x
- De Guzman, J. A., Thilakarathna, K., and Seneviratne, A. (2019). Security and Privacy Approaches in Mixed Reality: A Literature Survey. *ACM Comput. Surv. (Csur)* 52, 1–37.
- Dey, A. K., and Abowd, G. D. (2000). Towards a Better Understanding of Context and Context-Awareness. In Proceedings of the SIGCHI conference on Human factors in computing systems.
- Dogan, M. D., Acevedo Colon, S. V., Sinha, V., Akşit, K., and Mueller, S. (2021). “SensiCut: Material-Aware Laser Cutting Using Speckle Sensing and Deep Learning,” in Proceedings of the 34th Annual ACM Symposium on User Interface Software and Technology, Virtual Event USA (New York, NY: ACM). doi:10.1145/3472749.3474733
- Dogan, M. D., Faruqi, F., Churchill, A. D., Friedman, K., Cheng, L., Subramanian, S., et al. (2020). “G-id: Identifying 3d Prints Using Slicing Parameters,” in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–13. doi:10.1145/3313831.3376202
- Dourish, P., Adler, A., Bellotti, V., and Henderson, A. (1996). Your Place or Mine? Learning from Long-Term Use of Audio-Video Communication. *Comput. Supported Coop. Work* 5, 33–62. doi:10.1007/bf00141935
- Duchowski, A., Krejtz, K., Zurawska, J., and House, D. (2019). Using Microsaccades to Estimate Task Difficulty during Visual Search of Layered Surfaces. *IEEE Trans. visualization Comput. graphics*.
- Dun, X., Ikoma, H., Wetzstein, G., Wang, Z., Cheng, X., and Peng, Y. (2020). Learned Rotationally Symmetric Diffractive Achromat for Full-Spectrum Computational Imaging. *Optica* 7, 913–922. doi:10.1364/optica.394413
- Fares, R., Fang, S., and Komogortsev, O. (2013). *Can We Beat the Mouse with MAGIC?* New York, NY, USA: Association for Computing Machinery, 1387–1390.
- Fu, R., Wang, H., and Zhao, W. (2016). Dynamic Driver Fatigue Detection Using Hidden Markov Model in Real Driving Condition. *Expert Syst. Appl.* 63, 397–411. doi:10.1016/j.eswa.2016.06.042
- Fuchsberger, V. (2019). The Future’s Hybrid Nature. *Interactions* 26, 26–31. doi:10.1145/3328481
- Fuller, A., Fan, Z., Day, C., and Barlow, C. (2020). Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access* 8, 108952–108971. doi:10.1109/access.2020.2998358
- García, K., Mendoza, S., Decouchant, D., and Brézillon, P. (2018). Facilitating Resource Sharing and Selection in Ubiquitous Multi-User Environments. *Inf. Syst. Front.* 20, 1075–1095. doi:10.1007/s10796-016-9708-0
- Gardony, A. L., Lindeman, R. W., and Brunyé, T. T. (2020). “Eye-tracking for Human-Centered Mixed Reality: Promises and Challenges,” in *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality*. Editors B. C. Kress and C. Peroz (New York, NY: International Society for Optics and Photonics SPIE), 11310, 230–247. doi:10.1117/12.2542699
- Gebhardt, C., Hecox, B., van Opheusden, B., Wigdor, D., Hillis, J., Hilliges, O., et al. (2019). “Learning Cooperative Personalized Policies from Gaze Data,” in Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, 197–208. doi:10.1145/3332165.3347933
- Gee, J. P. (2003). What Video Games Have to Teach Us about Learning and Literacy. *Comput. Entertain.* 1, 20. doi:10.1145/950566.950595
- Glaessgen, E., and Stargel, D. (2012). The Digital Twin Paradigm for Future NASA and U.S. Air Force Vehicles. *NTRS - NASA*. doi:10.2514/6.2012-1818
- Grasset, R., Looser, J., and Billinghurst, M. (2006). “Transitional Interface: Concept, Issues and Framework,” in Proceedings of the 2006 IEEE/ACM International Symposium on Mixed and Augmented Reality (New York, NY: ISMAR), 231–232. doi:10.1109/ismar.2006.297819
- Grubert, J., Langlotz, T., Zollmann, S., and Regenbrecht, H. (2017). Towards Pervasive Augmented Reality: Context-Awareness in Augmented Reality. *IEEE Trans. Vis. Comput. Graphics* 23, 1706–1724. doi:10.1109/tvcg.2016.2543720
- Hamilton, D., McKechnie, J., Edgerton, E., and Wilson, C. (2021). Immersive Virtual Reality as a Pedagogical Tool in Education: a Systematic Literature Review of Quantitative Learning Outcomes and Experimental Design. *J. Comput. Educ.* 8, 1–32. doi:10.1007/s40692-020-00169-2
- Happa, J., Glencross, M., and Steed, A. (2019). Cyber Security Threats and Challenges in Collaborative Mixed-Reality. *Front. ICT* 6, 5. doi:10.3389/fict.2019.00005
- Hartmann, D., and Van der Auweraer, H. (2020). Digital Twins. arXiv preprint arXiv:2001.09747
- Hassib, M., Khamis, M., Friedl, S., Schneegass, S., and Alt, F. (2017a). “Brainatwork: Logging Cognitive Engagement and Tasks in the Workplace Using Electroencephalography,” in Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia, New York, NY, USA (New York, NY: Association for Computing Machinery), 305–310.
- Hassib, M., Schneegass, S., Eiglsperger, P., Henze, N., Schmidt, A., and Alt, F. (2017b). “Engagemeter: A System for Implicit Audience Engagement Sensing Using Electroencephalography,” in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, New York, NY, USA (New York, NY: Association for Computing Machinery), 5114–5119.
- Heimo, O. I., Kimppa, K. K., Helle, S., Korkalainen, T., and Lehtonen, T. (2014). “Augmented Reality - towards an Ethical Fantasy,” in 2014 IEEE International Symposium on Ethics in Science, Technology and Engineering, 1–7.
- Held, R. M., and Durlach, N. I. (1992). Telepresence. *Presence: Teleoperators & Virtual Environments* 1, 109–112. doi:10.1162/pres.1992.1.1.109
- Hernandez, J., Riobo, I., Rozga, A., Abowd, G. D., and Picard, R. W. (2014). “Using Electrodermal Activity to Recognize Ease of Engagement in Children during Social Interactions,” in Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (New York, NY: Association for Computing Machinery), 307–317. doi:10.1145/2632048.2636065
- Herz, M., and Rauschnabel, P. A. (2019). Understanding the Diffusion of Virtual Reality Glasses: The Role of media, Fashion and Technology. *Technol. Forecast. Soc. Change* 138, 228–242. doi:10.1016/j.techfore.2018.09.008
- Howcroft, J., Klejman, S., Fehlings, D., Wright, V., Zabjek, K., Andrysek, J., et al. (2012). Active Video Game Play in Children with Cerebral Palsy: Potential for Physical Activity Promotion and Rehabilitation Therapies. *Arch. Phys. Med. Rehabil.* 93, 1448–1456. doi:10.1016/j.apmr.2012.02.033
- Hutt, S., Mills, C., Bosch, N., Krasich, K., Brockmole, J., and D’Mello, S. (2017). “out of the Fr-Eye-Ing pan”: Towards Gaze-Based Models of Attention during Learning with Technology in the Classroom,” in Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization (Association for Computing Machinery) (New York, NY: UMAP), 94–103.
- Huws, U., Robinson, W. B., and Robinson, S. (1990). *Telework towards the Elusive Office*. John Wiley & Sons.
- Itoh, Y., Kaminokado, T., and Aksit, K. (2021). Beaming Displays. *IEEE Trans. Vis. Comput. Graphics* 27, 2659–2668. doi:10.1109/tvcg.2021.3067764
- Jain, D., Sra, M., Guo, J., Marques, R., Wu, R., Chiu, J., et al. (2016). “Immersive Terrestrial Scuba Diving Using Virtual Reality,” in Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, 1563–1569. doi:10.1145/2851581.2892503
- Kaviyaraj, R., and Uma, M. (2021). “A Survey on Future of Augmented Reality with Ai in Education,” in 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS IEEE), 47–52. doi:10.1109/icaais50930.2021.9395838
- Kelly, B. E., Bhattacharya, I., Heidari, H., Shusteff, M., Spadaccini, C. M., and Taylor, H. K. (2019). Volumetric Additive Manufacturing via Tomographic Reconstruction. *Science* 363, 1075–1079. doi:10.1126/science.aau7114

- Kim, J., Jeong, Y., Stengel, M., Akşit, K., Albert, R., Boudaoud, B., et al. (2019). Foveated AR. *ACM Trans. Graph.* 38, 1–15. doi:10.1145/3306346.3322987
- Kim, J., Zheng, C., Takahashi, H., Gross, M. D., Ashbrook, D., and Yeh, T. (2018). “Compositional 3d Printing: Expanding & Supporting Workflows towards Continuous Fabrication,” in Proceedings of the 2Nd ACM Symposium on Computational Fabrication, New York, NY, USA (New York, NY: ACM), 1–10.
- Kiyokawa, K., Billinghurst, M., Campbell, B., and Woods, E. (2003). “An Occlusion Capable Optical See-Through Head Mount Display for Supporting Co-located Collaboration,” in The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, Proceedings (IEEE), 133–141.
- Knierim, P., Wozniak, P. W., Abdelrahman, Y., and Schmidt, A. (2019). “Exploring the Potential of Augmented Reality in Domestic Environments,” in Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services, 1–12. doi:10.1145/3338286.3340142
- Kosmyna, N., and Maes, P. (2019). Attentivu: An Eeg-Based Closed-Loop Biofeedback System for Real-Time Monitoring and Improvement of Engagement for Personalized Learning. *Sensors (Basel)* 19. doi:10.3390/s19235200
- Kosmyna, N., Morris, C., Nguyen, T., Zepf, S., Hernandez, J., and Maes, P. (2019a). “Attentivu: Designing Eeg and Eog Compatible Glasses for Physiological Sensing and Feedback in the Car,” in Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, New York, NY, USA (New York, NY: Association for Computing Machinery), 355–368.
- Kosmyna, N., Morris, C., Sarawagi, U., Nguyen, T., and Maes, P. (2019b2019). “Attentivu: A Wearable Pair of Eeg and Eog Glasses for Real-Time Physiological Processing,” in IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN), 1–4. doi:10.1109/bsn.2019.8771080
- Kosmyna, N., Tarpin-Bernard, F., and Rivet, B. (2015). “Towards Brain Computer Interfaces for Recreational Activities: Piloting a Drone,” in *Human-Computer Interaction – INTERACT 2015*. Editors J. Abascal, S. Barbosa, M. Fetter, T. Gross, P. Palanque, and M. Winckler (Cham: Springer International Publishing), 506–522. doi:10.1007/978-3-319-22701-6_37
- Koulieris, G. A., Akşit, K., Stengel, M., Mantiuk, R. K., Mania, K., and Richardt, C. (2019). Near-Eye Display and Tracking Technologies for Virtual and Augmented Reality. *Computer Graphics Forum* 38, 493–519. doi:10.1111/cgf.13654
- Kranzberg, M. (1986). Technology and History: “Kranzberg’s Laws”. *Technology Cult.* 27, 544–560. doi:10.2307/3105385
- Krogmeier, C., Mousas, C., and Whittinghill, D. (2019). Human-virtual Character Interaction: Toward Understanding the Influence of Haptic Feedback. *Computer Animation Virtual Worlds* 30, e1883. doi:10.1002/cav.1883
- Kuo, G., Waller, L., Ng, R., and Maimone, A. (2020). High Resolution Étendue Expansion for Holographic Displays. *ACM Trans. Graphics (Tog)* 66, 1–66. doi:10.1145/3386569.3392414
- Kwak, H., Blackburn, J., and Han, S. (2015). “Exploring Cyberbullying and Other Toxic Behavior in Team Competition Online Games,” in Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, New York, NY, USA (New York, NY: ACM), 3739–3748. doi:10.1145/2702123.2702529
- Langlotz, T., Sutton, J., Zollmann, S., and Itoh, H. (2018). “Chromaglasses: Computational Glasses for Compensating Colour Blindness,” in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, New York, NY, USA (New York, NY: ACM), 112–390.
- Leroy, S. (2009). Why Is it So Hard to Do My Work? the challenge of Attention Residue when Switching between Work Tasks. *Organizational Behav. Hum. Decis. Process.* 109, 168–181. doi:10.1016/j.obhdp.2009.04.002
- Li, R., Whitmire, E., Stengel, M., Boudaoud, B., Kautz, J., Luebke, D., et al. (2020). Optical Gaze Tracking with Spatially-Sparse Single-Pixel Detectors. arXiv preprint arXiv:2009.06875
- Li, T., Liu, Q., and Zhou, X. (2017). “Ultra-low Power Gaze Tracking for Virtual Reality,” in Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems, 114. doi:10.1145/3131672.3131682
- Lincoln, P., Blate, A., Singh, M., Whitted, T., State, A., Lastra, A., et al. (2016). From Motion to Photons in 80 Microseconds: Towards Minimal Latency for Virtual and Augmented Reality. *IEEE Trans. Vis. Comput. Graphics* 22, 1367–1376. doi:10.1109/tvcg.2016.2518038
- Liu, C., Plopski, A., and Orlosky, J. (2020). Orthogaze: Gaze-Based Three-Dimensional Object Manipulation Using Orthogonal Planes. *Comput. Graphics*. doi:10.1016/j.cag.2020.04.005
- Liu, Z., Meyendorf, N., and Mrad, N. (2018). The Role of Data Fusion in Predictive Maintenance Using Digital Twin. *AIP Conf. Proc.* 1949, 020023. doi:10.1063/1.5031520
- Lu, C., Chakravarthula, P., Tao, Y., Chen, S., and Fuchs, H. (2020). “Improved Vergence and Accommodation via Purkinje Image Tracking with Multiple Cameras for Ar Glasses,” in 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (IEEE). doi:10.1109/ismar50242.2020.00058
- Mack, K., McDonnell, E., Jain, D., Lu Wang, L., Froehlich, E. J., and Findlater, L. (2021). “What Do We Mean by ‘Accessibility Research’? a Literature Survey of Accessibility Papers in Chi and Assets from 1994 to 2019,” in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 1–18.
- Mahar, K., Zhang, A. X., and Karger, D. (2018a). “Squadbox: A Tool to Combat Email Harassment Using Friendsourced Moderation,” in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, New York, NY, USA (New York, NY: ACM), 1586–5863.
- Mahar, K., Zhang, A. X., and Karger, D. (2018b). “Squadbox: A Tool to Combat Email Harassment Using Friendsourced Moderation,” in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 1–13.
- Marwecki, S., Wilson, A. D., Ofek, E., Gonzalez Franco, M., and Holz, C. (2019). “Mise-unseen: Using Eye Tracking to Hide Virtual Reality Scene Changes in plain Sight,” in Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, 777–789.
- Matsukura, H., Yoneda, T., and Ishida, H. (2013). Smelling Screen: Development and Evaluation of an Olfactory Display System for Presenting a Virtual Odor Source. *IEEE Trans. Vis. Comput. Graphics* 19, 606–615. doi:10.1109/tvcg.2013.40
- Mayer, S., Ciortea, A., Ricci, A., Robles, M. I., Kovatsch, M., and Croatti, A. (2018). Hypermedia to Connect Them All: Autonomous Hypermedia Agents and Socio-Technical Interactions. *Internet Technology Lett.* 1, e50. doi:10.1002/itl2.50
- Milgram, P., Takemura, H., Utsumi, A., and Kishino, F. (1995). Augmented Reality: A Class of Displays on the Reality-Virtuality Continuum, *Telematicoperator telepresence Tech.*, 2351. 282–292.
- Miller, M. R., Herrera, F., Jun, H., Landay, J. A., and Bailenson, J. N. (2020). Personal Identifiability of User Tracking Data during Observation of 360-degree VR Video. *Sci. Rep.* 10, 17404. doi:10.1038/s41598-020-74486-y
- Miyashita, H. (2020). “Norimaki Synthesizer: Taste Display Using Ion Electrophoresis in Five Gels,” in Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, 1–6. doi:10.1145/3334480.3382984
- Montgomery, L. D., Montgomery, R. W., and Guisado, R. (1995). Rheoencephalographic and Electroencephalographic Measures of Cognitive Workload: Analytical Procedures. *Biol. Psychol.* 40, 143–159. doi:10.1016/0301-0511(95)05117-1
- Mott, M., Tang, J., Kane, S., Cutrell, E., and Ringel Morris, M. (2020). “I Just Went into it Assuming that I Wouldn’t Be Able to Have the Full Experience: Understanding the Accessibility of Virtual Reality for People with Limited Mobility,” in The 22nd International ACM SIGACCESS Conference on Computers and Accessibility, New York, NY, USA (New York, NY: Association for Computing Machinery). doi:10.1145/3373625.3416998
- Mueller, F. F., Lopes, P., Strohmeier, P., Ju, W., Seim, C., Weigel, M., et al. (2020). “Next Steps for Human-Computer Integration,” in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–15. doi:10.1145/3313831.3376242
- Mueller, S., Fritzsche, M., Kossmann, J., Schneider, M., Striebel, J., and Baudisch, P. (2015). “Scotty: Relocating Physical Objects across Distances Using Destructive Scanning, Encryption, and 3d Printing,” in Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction, New York, NY, USA (New York, NY: Association for Computing Machinery), 233–240.
- Nakano, K., Horita, D., Sakata, N., Kiyokawa, K., Yanai, K., and Narumi, T. (2019). “Deeptaste: Augmented Reality Gustatory Manipulation with gan-based Real-Time Food-To-Food Translation,” in Proceedings of the 2019 IEEE

- International Symposium on Mixed and Augmented Reality (New York, NY: ISMAR), 328–339. doi:10.1109/ismar.2019.000-1
- Narumi, T., Kajinami, T., Nishizaka, S., Tanikawa, T., and Hirose, M. (2011a/2011). “Pseudo-gustatory Display System Based on Cross-Modal Integration of Vision, Olfaction and Gustation,” in IEEE Virtual Reality Conference (IEEE), 127–130. doi:10.1109/vr.2011.5759450
- Narumi, T., Nishizaka, S., Kajinami, T., Tanikawa, T., and Hirose, M. (2011b). “Meta Cookie+: an Illusion-Based Gustatory Display,” in International Conference on Virtual and Mixed Reality (Springer), 260–269. doi:10.1007/978-3-642-22021-0_29
- Norouzi, N., Bruder, G., Belna, B., Mutter, S., Turgut, D., and Welch, G. (2019). A Systematic Review of the Convergence of Augmented Reality, Intelligent Virtual Agents, and the Internet of Things. *Artif. intelligence IoT*, 1–24. doi:10.1007/978-3-030-04110-6_1
- Orts-Escolano, S., Rhemann, C., Fanello, S., Chang, W., Kowdle, A., Degtyarev, Y., et al. (2016). “Holoportation: Virtual 3d Teleportation in Real-Time,” in Proceedings of the 29th Annual Symposium on User Interface Software and Technology, New York, NY, USA (New York, NY: Association for Computing Machinery), 741–754. doi:10.1145/2984511.2984517
- Pellas, N., Kazanidis, I., and Palaigeorgiou, G. (2020). A Systematic Literature Review of Mixed Reality Environments in K-12 Education. *Educ. Inf. Technol.* 25, 2481–2520. doi:10.1007/s10639-019-10076-4
- Peng, H., Briggs, J., Wang, C.-Y., Guo, K., Kider, J., Mueller, S., et al. (2018). “Roma: Interactive Fabrication with Augmented Reality and a Robotic 3d Printer,” in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, New York, NY, USA (New York, NY: Association for Computing Machinery), 1–12.
- Peng, H., Wu, R., Marschner, S., and Guimbretière, F. (2016). “On-the-fly Print: Incremental Printing while Modelling,” in Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 887–896.
- Peng, Y., Choi, S., Padmanaban, N., and Wetzstein, G. (2020). Neural Holography with Camera-In-The-Loop Training. *ACM Trans. Graph.* 39, 1–14. doi:10.1145/3414685.3417802
- Peng, Y., Sun, Q., Dun, X., Wetzstein, G., Heidrich, W., and Heide, F. (2019). Learned Large Field-Of-View Imaging with Thin-Plate Optics. *ACM Trans. Graph.* 38, 1–14. doi:10.1145/3355089.3356526
- Piovarci, M., Wessely, M., Jagielski, M., Alexa, M., Matusik, W., and Didyk, P. (2017). “Directional Screens,” in Proceedings of the 1st Annual ACM Symposium on Computational Fabrication, 1–10.
- Popescu, V. G., Burdea, G. C., Bouzit, M., and Hentz, V. R. (2000). A Virtual-Reality-Based Telerehabilitation System with Force Feedback. *IEEE Trans. Inform. Technol. Biomed.* 4, 45–51. doi:10.1109/4233.826858
- Raca, M., Kidzinski, L., and Dillenbourg, P. (2015). “Translating Head Motion into Attention - towards Processing of Student’s Body-Language,” in EDM Proceedings of the 8th International Conference on Educational Data Mining, Madrid, Spain, June 26–29, 2015.
- Ranasinghe, N., Jain, P., Karwita, S., Tolley, D., and Do, E. Y.-L. (2017). “Ambiotherm: Enhancing Sense of Presence in Virtual Reality by Simulating Real-World Environmental Conditions,” in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 1731–1742.
- Rathinavel, K., Wetzstein, G., and Fuchs, H. (2019). Varifocal Occlusion-Capable Optical See-Through Augmented Reality Display Based on Focus-Tunable Optics. *IEEE Trans. Vis. Comput. Graphics* 25, 3125–3134. doi:10.1109/tvcg.2019.2933120
- Rauschnabel, P. A., Brem, A., and Ivens, B. S. (2015). Who Will Buy Smart Glasses? Empirical Results of Two Pre-market-entry Studies on the Role of Personality in Individual Awareness and Intended Adoption of Google Glass Wearables. *Comput. Hum. Behav.* 49, 635–647. doi:10.1016/j.chb.2015.03.003
- Rauschnabel, P. A., and Ro, Y. K. (2016). Augmented Reality Smart Glasses: An Investigation of Technology Acceptance Drivers. *Ijtmkt* 11, 123–148. doi:10.1504/ijtmkt.2016.075690
- Ricci, A., Piunti, M., Tummolini, L., and Castelfranchi, C. (2015). The Mirror World: Preparing for Mixed-Reality Living. *IEEE Pervasive Comput.* 14, 60–63. doi:10.1109/mpvr.2015.44
- Risteska Stojkoska, B. L., and Trivodaliev, K. V. (2017). A Review of Internet of Things for Smart home: Challenges and Solutions. *J. Clean. Prod.* 140, 1454–1464. doi:10.1016/j.jclepro.2016.10.006
- Rizzo, A. S., Difede, J., Rothbaum, B. O., Reger, G., Spitalnick, J., Cukor, J., et al. (2010). Development and Early Evaluation of the Virtual Iraq/Afghanistan Exposure Therapy System for Combat-Related PTSD. *Ann. New York Acad. Sci.* 1208, 114–125. doi:10.1111/j.1749-6632.2010.05755.x
- Roy, A. K., Akhtar, M. N., Mahadevappa, M., Guha, R., and Mukherjee, J. (2017). A Novel Technique to Develop Cognitive Models for Ambiguous Image Identification Using Eye Tracker. *IEEE Trans. Affective Comput.* 11, 63–77.
- Samad, M., Gatti, E., Hermes, A., Benko, H., and Parise, C. (2019). “Pseudo-haptic Weight: Changing the Perceived Weight of Virtual Objects by Manipulating Control-Display Ratio,” in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–13.
- Sato, W., Fujimura, T., and Suzuki, N. (2008). Enhanced Facial EMG Activity in Response to Dynamic Facial Expressions. *Int. J. Psychophysiology* 70, 70–74. doi:10.1016/j.jpsycho.2008.06.001
- Schäfer, A., Reis, G., and Stricker, D. (2021). A Survey on Synchronous Augmented, Virtual and Mixed Reality Remote Collaboration Systems. arXiv preprint arXiv:2102.05998
- Sereno, M., Wang, X., Besançon, L., McGuffin, M. J., and Isenberg, T. (2020). Collaborative Work in Augmented Reality: A Survey. *IEEE Trans. Visualization Computer Graphics*. doi:10.1109/tvcg.2020.3032761
- Shneiderman, B. (2020). Human-centered Artificial Intelligence: Three Fresh Ideas. *AIS Trans. Human-Computer Interaction* 12, 109–124. doi:10.17705/1thci.00131
- Sohlberg, M. M., and Mateer, C. A. (1987). Effectiveness of an Attention-Training Program. *J. Clin. Exp. Neuropsychol.* 9, 117–130. doi:10.1080/01688638708405352
- Sra, M., Xu, X., Mottelson, A., and Maes, P. (2018). “Vmotion: Designing a Seamless Walking Experience in Vr,” in Proceedings of the 2018 Designing Interactive Systems Conference, 59–70.
- Steed, A., and Oliveira, M. F. (2009). *Networked Graphics: Building Networked Games and Virtual Environments*. USA: Elsevier.
- Steed, A., Ortega, F. R., Williams, A. S., Kruijff, E., Stuerzlinger, W., Batmaz, A. U., et al. (2020). Evaluating Immersive Experiences during Covid-19 and beyond. *Interactions* 27, 62–67. doi:10.1145/3406098
- Stengel, M., Grogoric, S., Eisemann, M., Eisemann, E., and Magnor, M. A. (2015). “An Affordable Solution for Binocular Eye Tracking and Calibration in Head-Mounted Displays,” in Proceedings of the 23rd ACM international conference on Multimedia, 15–24. doi:10.1145/2733373.2806265
- Stengel, M., and Magnor, M. (2016). Gaze-contingent Computational Displays: Boosting Perceptual Fidelity. *IEEE Signal. Process. Mag.* 33, 139–148. doi:10.1109/msp.2016.2580913
- Sutherland, I. (1965). *The Ultimate Display*. Göteborg, Sweden: CHALMERS UNIVERSITY OF TECHNOLOGY.
- Sweller, J., Ayres, P., and Kalyuga, S. (2011). *Measuring Cognitive Load*. New York, NY: Springer New York, 71–85. doi:10.1007/978-1-4419-8126-4_6
- Tao, F., Zhang, H., Liu, A., and Nee, A. Y. C. (2019). Digital Twin in Industry: State-Of-The-Art. *IEEE Trans. Ind. Inf.* 15, 2405–2415. doi:10.1109/tii.2018.2873186
- Templier, T., Bektas, K., and Hahnloser, R. H. (2016). *Eye-Trace: Segmentation of Volumetric Microscopy Images with Eyegaze*. New York, NY, USA: Association for Computing Machinery, 5812–5823.
- Topal, C., Gunal, S., Koçdeviren, O., Doğan, A., and Gerek, Ö. N. (2013). A Low-Computational Approach on Gaze Estimation with Eye Touch System. *IEEE Trans. Cybern.* 44, 228–239. doi:10.1109/TCYB.2013.2252792
- Urey, H., Chellappan, K. V., Erden, E., and Surman, P. (2011). State of the Art in Stereoscopic and Autostereoscopic Displays. *Proc. IEEE* 99, 540–555. doi:10.1109/jproc.2010.2098351
- van Dokkum, L. E. H., Ward, T., and Laffont, I. (2015). Brain Computer Interfaces for Neurorehabilitation - its Current Status as a Rehabilitation Strategy post-stroke. *Ann. Phys. Rehabil. Med.* 58, 3–8. doi:10.1016/j.rehab.2014.09.016
- Wang, G. G. (2002). Definition and Review of Virtual Prototyping. *J. Comput. Inf. Sci. Eng.* 2, 232–236. doi:10.1115/1.1526508
- Weiser, M. (1999). The Computer for the 21st Century. *SIGMOBILE Mob. Comput. Commun. Rev.* 3, 3–11. doi:10.1145/329124.329126
- Whitmire, E., Trutoiu, L., Cavin, R., Perck, D., Scally, B., Phillips, J., et al. (2016). “Eyecontact: Scleral Coil Eye Tracking for Virtual Reality,” in Proceedings of the 2016 ACM International Symposium on Wearable Computers, 184–191.

- Yarramreddy, A., Gromkowski, P., and Baggili, I. (2018). "Forensic Analysis of Immersive Virtual Reality Social Applications: A Primary Account," in 2018 IEEE Security and Privacy Workshops (New York, NY: SPW), 186–196. doi:10.1109/spw.2018.00034
- Zander, T. O., Kothe, C., Jatzev, S., and Gaertner, M. (2010). "Enhancing Human-Computer Interaction with Input from Active and Passive Brain-Computer Interfaces," in *Brain-Computer Interfaces*. doi:10.1007/978-1-84996-272-8_11
- Zotter, F., and Frank, M. (2019). *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*. Springer Nature.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Orlosky, Sra, Bektaş, Peng, Kim, Kos'myna, Höllerer, Steed, Kiyokawa and Akşit. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.