Invited Paper

Virtual and Augmented Reality on the 5G Highway

Jason Orlosky[1,a]   Kiyoshi Kiyokawa[1]   Haruo Takemura[1]

Abstract: In recent years, virtual and augmented reality have begun to take advantage of the high speed capabilities of data streaming technologies and wireless networks. However, limitations like bandwidth and latency still prevent us from achieving high fidelity telepresence and collaborative virtual and augmented reality applications. Fortunately, both researchers and engineers are aware of these problems and have set out to design 5G networks to help us to move to the next generation of virtual interfaces. This paper reviews state of the art virtual and augmented reality communications technology and outlines current efforts to design an effective, ubiquitous 5G network to help to adapt to virtual application demands. We discuss application needs in domains like telepresence, education, healthcare, streaming media, and haptics, and provide guidelines and future directions for growth based on this new network infrastructure.

Keywords: virtual reality, augmented reality, mixed reality, 5G Network

1. Introduction

One staple of effective electronic communication has always been a robust, available network connection. From the birth of the Internet to the advent of modern high bandwidth cellphone networks, we have strived to provide faster services with lower latency and broader coverage. We are now on the brink of implementing a faster than ever 5G network infrastructure, which has significant implications for the future success of virtual and augmented realities (AR and VR). Researchers must also begin thinking about the implications of such a network and its relationship to virtual technologies, as factors like security, privacy, and safety will be of utmost importance. In the following, we first delve into the history and most promising applications of VR and AR, provide an overview of important state-of-the-art research, and discuss applications and future possibilities that 5G can help enable.

1.1 A Brief History of Virtual Reality

Virtual reality is a technology that immerses a user into a computer generated virtual world by replacing the real world around the user with a virtual one. With the ultimate virtual reality system, a user would not be able to distinguish the virtual world from the real. This idea, embodied by a pair of goggles that let the user experience a fictional world through five senses, was first presented in a science fiction called Pygmalion's Spectacles by Weinbaum in the 1930s [59]. Shortly after in the 1950s, Heilig developed a multi-modal theater cabinet called Sensorama that displayed stereoscopic 3D images, stereo sound, aromas, winds and vibrations during film [20]. In 1968, Sutherland demonstrated the first functioning, see-through head mounted display (HMD) with head tracking capability [51]. The user was able

to see computer generated 3D wireframe objects overlaid on the real environment from his or her own perspective in real time. This system is considered to be the father of both virtual and augmented reality today.

One of the first large scale networked VR systems was SIM-NET, a military simulation for DARPA, developed in 1983 [55], as shown in **Fig. 1**. In late 80s, the Naval Postgraduate School developed NPSNET, a battlefield simulation system supporting hundreds of users at a time [62]. Other examples of networked VR systems in early days include DIVE [10] and MASSIVE [18]. In 1987, Jaron Lanier coined the term 'virtual reality' and launched a company called VPL Research. One of its first products, the EyePhone, also considered the first commercial head mounted display, was then introduced in many research laboratories, which contributed to the first boom of virtual reality in early 1990s. Today's virtual reality boom is thought to have been triggered by inexpensive smartphone based HMDs. One of the first such systems was the FOV2GO papercraft HMD developed at University of Southern California in 2012 [14]. Originating from the same research group, the Oculus Rift was also launched in the same year. According to SuperData Research, Statista 2016, the virtual reality market has a value of approximately $3.7B worldwide as of 2016, and is expected to grow to $40B in 2020. To date, however, most networked virtual reality systems only exchange event triggers that change shared content such as user motion and key presses, so that the required network traffic is minimal. As we will describe later, the 5G high speed network will allow real time large data transmission such as HD 3D video, high fidelity virtual environments, and other streamed VR experiences for networked displays and viewing systems.

1.2 A Brief History of Augmented Reality

Augmented reality (AR) is a technology that overlays computer generated information onto the real world. Thus, our environment is 'augmented' so that the user can perform the task

---

[1]   Cybermedia Center, Osaka University, Toyonaka, Osaka 560–0043, Japan
[a]   orlosky@lab.ime.cmc.osaka-u.ac.jp

Fig. 1   Images showing early implementations of SIMNET (Courtesy of Jim Millar [37]).

at hand with less effort. As such, the goals of augmented reality have traditionally been task driven. One objective of the ultimate augmented reality system would be show the most relevant information at the most appropriate time and place. Another type of ultimate augmented reality system would add virtual content onto the real world or modify existing objects in a way that is indistinguishable to the user. As mentioned above, this basic idea was presented and demonstrated by Sutherland in the1960s. Somewhat later, Caudell and Mizell coined the term 'augmented reality' in 1992 for the HMD-based system they developed for wire bundle assembly at Boeing [12]. Released in 1999, the marker-based open source tracking library, ARToolKit, further facilitated the acceleration and dissemination of augmented reality [29]. In the late 2000s, augmented reality was made known to general public through the entertainment industry using games such as The Eye of Judgement in 2007, which is thought to be the first AR consumer game, and the Nintendo 3DS in 2011. To date, Pokemon Go alone has acquired more than 200 million players and earns more than $2M everyday worldwide. According to Digi-Capital, the augmented reality market as a whole is expected to grow to $90B by 2020.

Technically speaking, virtual and augmented reality systems are quite similar to each other. Both require sensing (input) and display (output) subsystems as well as a scene management subsystem. Sensing subsystems need to keep track of user position (and motion) and accept some form of command input. Display subsystems render a 3D scene according to user position, and output other sensory information like sound. Hardware and software architecture for the two systems are also quite similar or almost identical. For both, minimizing end-to-end (motion to photon) latency is essential since latency degrades user comfort in VR and visual quality in AR, and has a number of other negative effects. Accurate position tracking (or registration in augmented reality) is also important for both systems, but is typically more difficult to achieve for AR because of the mobile nature of the applications. In general, display hardware for AR is also more difficult to develop than for virtual reality, due to the complex nature of see through optics and the relationship between virtual content and the environment. Supplemental material is typically drawn in AR, so comparable rendering requirements are often higher for VR since the whole scene must be drawn every frame, and often at higher frame rates.

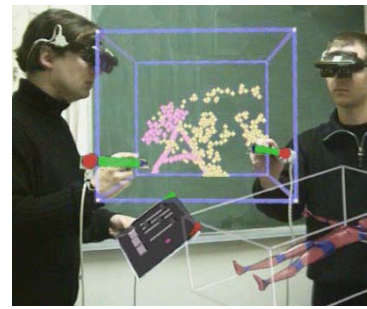In addition to standalone virtual technologies, early examples

Fig. 2   Typical early collaborative AR systems. (top) Studierstube (Courtesy of Dieter Schmalstieg [47]), (middle) Shared Space (Courtesy of Mark Billinghurst [7]), (bottom) SeamlessDesign [31].

of networked and/or collaborative AR systems started to come about in the early nineties, such as Studierstube [47] (**Fig. 2**, top) and Shared Space [7] (Fig. 2, middle). Such co-located AR systems are beneficial for intensive collaborative work such as 3D modeling (Fig. 2, bottom, [31]) since discussions and exchanging ideas between participants can happen more naturally and smoothly due to non-verbal communication cues present while sharing and viewing virtual content [32].

These developments allowed for the sharing of virtual content between users from different perspectives, but bandwidth requirements were minimal for demonstrating the basic technology. Some other networked AR systems target remote collaboration or remote assistance with real time video transmission [8], but the bandwidth and latency requirements of today's virtual applications are much more demanding than those of research projects twenty years ago. Using the 5G high speed network, current and future networked AR systems will have richer content and facilitate more effective collaboration on the go, as we will discuss later.

Next, we describe communications challenges, the 5G network, and recent, relevant technology developments. Then, we discuss how these developments will affect AR and VR interfaces, help solve existing problems, and enable new content and interactions. We conclude with directions for the near and distant future of AR and VR as a whole. In short, our goal is to outline

how the 5G network will influence the field, and help provide both an overview and guidelines for researchers, engineers, and businesses to help successfully adapt to these changes in infrastructure.

2. State of the Art AR/VR Technologies

With the progress of the high-speed mobile network and the increasing number of smart devices, the global digital content market is expected to expand from $475B in 2018 to $1174B in 2025. Virtual and augmented realities are considered key technologies to take the initiative in such a future digital content market, with just the market for VR/AR content expected to grow from $1934M in 2018 to $3264M in 2025 [38].

2.1 Game Technology

The most recent example of augmented reality that has taken the world by storm is that of Pokemon Go. Though this game does not use the most complex AR or network technologies, its design facilitates an interactive, social, mobile game that has become extremely popular in a short period of time. We expect other similar games, especially those that begin to take advantage of more advanced wearable displays, to become prevalent and increasingly popular with the appearance of the 5G network. Much like the telepresence applications described later, 5G will allow players to interact with virtual content as a group or in a social manner that will make AR gaming even more attractive. 5G promises to solve many of the problems preventing such applications from running in real time today.

Though early examples of games like AR Quake [54] showed the potential for outdoor gaming, first person shooters, for example, require very low latency throughput in order to properly synchronize the positions and interactions of a large group of players. All players must have a similar low throughput for this to work effectively, or some players, if not all, will experience lag or dropouts. Because of how virtual content is dependent on the real world, extremely low latency (less than 40ms required, less than 8ms desired) is often a barrier to making mobile AR or VR gaming a reality when multiple players are present. We have seen that devices like the Microsoft Hololens allow single users to view content and play games within the confines of a building or room, primarily due to advances in inside-out tracking technology. To integrate other latency dependent entities into games, the 5G network is a significant step towards reaching these latency requirements [5].

2.2 Reconstruction

One fundamental technology essential to many AR and future VR applications is 3D reconstruction. This refers to the construction and localization of points, lines, or planes that represent real objects in 3D space. With such reconstructions, the ability to place virtual replicas of real objects in the real or virtual world becomes possible. Through the knowledge of the points representing the real world and virtual object, it becomes possible to recognize what parts of the 3D object should be visible and which should not, a concept referred to as occlusion. The important thing to realize is that these reconstructions, especially dense re-

Fig. 3 Image of Microsoft's Holoportation (Courtesy of Shahram Izadi [43]).

constructions that more accurately represent the real world, are often composed of millions of points, voxels, texels, or 3D structures. Though many systems have realized dense 3D reconstructions through techniques like DTAM [39] or KinectFusion [40], the constant sharing and synchronization of these environments between users is still a significant challenge.

2.3 Telepresence

Though AR and VR are already affecting many fields of research today, one application bringing a lot of excitement is that of Telepresence, also referred to as Telexistence. This concept refers to the ability to be present at a remote location in some form or fashion. The most rudimentary forms of telepresence date back hundreds of years, including the telephone. The amount of presence felt by both the local and remote users are limited by the technology used, where complete telepresence would be a 100% accurate representation of the two users so that they could not tell the difference between their real and virtual selves. Current technologies allow us to replicate this goal to a certain extent. For example, phones allow us to replicate the human voice, though with limited quality; Videoconferencing allows us to replicate a visual 2D representation of the human body; Most recently, methods utilizing volumetric reconstruction, like Holoportation [43] (**Fig. 3**) and Fusion 4D [13], have begun to scratch the surface of a 3D visual representation of the remote party; Simple physical reconstructions of objects or of a remote person's limbs has also recently become possible [26]. These technologies seek to create a complete auditory, visual, and haptic copy of individuals or items at a remote location, though limitations in technology, especially those related to wireless network throughput [1], are still preventing these applications from becoming a reality. Operating humanoid robots or avatars over the network [15] also suffer from these kinds of limitations.

2.4 Displays

The bandwidth requirements of virtual and augmented reality also depend on the pixel resolution and refresh rates of display hardware. For example, the first commercial HMD, the VPL EyePhone in 1989, had a fairly wide field of view, (75.3 × 58.4 degrees per eye, 90.0 degrees of binocular field of view) which is comparable to today's Oculus Rift and HTC Vive, but its resolution was very coarse. It had a monochrome array of 360 × 240 (86,400) cells with red, green and blue filters, resulting in
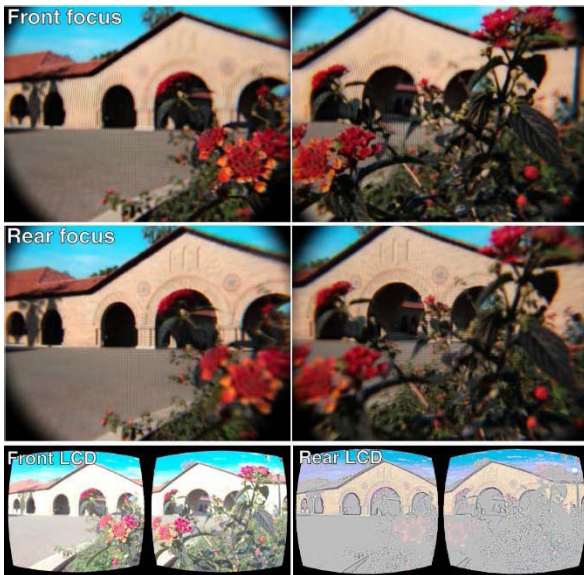
Fig. 4 Images of a Light-Field Stereoscope, showing resulting front and rear focus (top and middle rows, respectively), and the images rendered on the front and rear LCD to generate this multi-focal possibility. (Courtesy of Gordon Wetzstein [22]).

28,800 color triads or approximately 207.8×138.6 color pixels. The refresh rate was 30 Hz (it often dropped below 10 Hz due to low graphics power), and with 24 bits/pixel for stereo rendering, the uncompressed data rate was about 39.6 Mbps, or 2 Mbps in H.264 format. Modern HMDs have much higher resolution and refresh rates. For example, the resolution of Oculus Rift CV1 is 1080×1200 (1,296,000) per eye driven at 90 Hz resulting in an uncompressed data rate of about 5.3 Gbps, 135 times more than that of EyePhone. Even with H.264 format, this is around 250 Mbps, which would only be transmissible while mobile via the 5G network. Next generation technologies like light field and holographic displays reproduce more realistic 3D images with accommodation cues, and they demand even higher bandwidths. For example, light field displays with micro lens arrays typically require about 10×10 (100 times) the number of pixels compared to conventional binocular stereo displays. To achieve similar angular resolution, bandwidth of a few hundred Gbps is required. Though still bandwidth intensive, factored light field synthesis via stacked LCD panels, as shown in **Fig. 4** [22], require less data to produce the same resolution, which makes it a promising technology in the era of the 5G high speed network.

3. AR/VR in the Era of 5G Highway

3.1 Challenges
(1) Latency
Perhaps the most important and challenging aspect of achieving useful AR/VR systems is reducing delay (used interchangeably with latency in this paper). Many types of delay exist that can affect perception and performance, including sensor, network, and rendering delay. For the most part, sensor delay has been reduced to an amount that is imperceptible by humans in both AR devices like the HoloLens [17] and VR devices like the Oculus Rift or HTC Vive. The remaining problems arise when delay is introduced into the system that is not sensor-specific, such as sending 3D model data over the network in real time.

For individual AR and VR experiences, this latency is one-way. However, for games and haptic applications, this latency is bi-directional, effectively doubling the infrastructure requirements. Even with today's high speed networks, loading a dense point cloud model over a network of any significant distance is not feasible. Some applications have been able to alleviate this problem with clever rendering techniques, such as games that sending very basic positioning data over the network and let client systems handle rendering, though they haven't completely been able to solve the problem. Other applications like the recently developed Fusion 4D [13] involve much more complex models where latency can have a larger negative effect on results. Haptic interaction is also primarily latency intensive. Much like typing on a keyboard requires immediate haptic response for accurate text entry [21], haptic feedback can be perceived as unnatural or fake, if it is over 7 ms in inking applications for example [3], if the feedback is engaged too long after the visual stimulus has occurred.

(2) Bandwidth
Like latency, bandwidth is also a significant limitation for applications like 3D reconstruction and telepresence, where the transmission of gigabytes of dense model or point cloud data is a necessity. Pre-processing this information and just sending textures or frames to be rendered on a remote display is possible, but this still must be done at extremely low latency and bandwidth. "With 4G technologies, there are still barriers to provide services requiring both real-time response and big data sizes, which are expected to be a key requirement of future services. For example, the current 4G/LTE network is not able to provide instantaneous cloud services to mobile users which can enable people to have desktop-like experiences on the go as such cloud based services are expected to be much more commonly accessible by the public in societies of the 2020s [38]." Considering the new development of light field and holographic displays, sending direct data streams of video frames will have also extremely high bandwidth requirements.

(3) Quality of Service (QoS)
Even if we can solve the problems associated with latency and bandwidth, quality of service is a necessity for most applications. If applications like Skype, Facetime, or Google Hangouts suffer from frequent disconnects or degradation in image quality, users will turn to other applications. Similarly, if multiplayer virtual reality or applications for AR assistance in industrial environments suffer from lag spikes or dropouts, players and users will be reluctant to adopt these new technologies for use in the field.

(4) Availability
Availability, also highly tied to QoS, is a must if we are to foster worldwide adoption of AR and VR. Cell phones have made their way into the hands of over 90% of people on this planet, largely due to the availability of 3G and beyond networks distributed over vast areas of land.

(5) Colocation (i.e., separate physical spaces)
Colocation, or presence of different individuals in separate physical spaces, also poses problems to achieving meaningful AR and VR experiences. Because of the nature of this type of application, both latency and bandwidth play immense roles in accurately representing a remote environment. For example, when

two individuals in different rooms want to share a similar virtual space or objects, a myriad of reconstruction, real time capture, and haptic feedback mechanisms usually need to exist. End to end latency of reconstruction cameras, network, and rendering can often amount to seconds, which is unacceptable for many applications like tele-surgery.

(6) **Security and Safety**

Most use of AR and VR to date has been for inside or in-home use, which for the most part is relatively secure. Once mobile interfaces on wireless networks come into play, security will play a much more integral role, especially when business meetings, games, and private conversations can occur in virtual environments. "It is known that 2010 to 2020 will become The Decade of Ubiquity, which is defined as the next ten years where every aspect of our lives will be permeated by digital, mobile, media, data, information, augmented, virtual, and so forth [57]." For these applications, we will need to ensure that private information stays private, that users can appropriately and knowledgeably set application privileges and access levels with ease, and that mobile applications do not cause physical dangers to users due to distractions or limited perception of the real world.

(7) **Social Adoption**

As with any technology, social adoption of a new device or system usually takes time. Much like carrying a phone on your person was a rare occurrence 30 years ago, wearing an AR display like the Hololens while walking down the street would be met with awe. However, we have already seen the slow migration of VR into everyday life, despite initial resistance and ridicule of the technology in popular media. The same goes for hand-held AR systems like Pokemon Go. This adoption has not come without its problems though. We have seen incidences of pool players attempting to lean on virtual tables that do not exist (and falling), Pokemon Go players being robbed or beaten in parks while not paying attention to their surroundings, and some potentially health-hazardous instances of people that are terrified by VR simulations of roller coasters or horror games. As such, societal rules, safety mechanisms, and interactive devices that intelligently regulate presence and environmental perception need to be built into AR and VR interfaces for social adoption to be a success. Smart networks will also play a role in achieving safe and practical virtual systems by ensuring secure connections and environments for gaming and collaboration.

3.2 5G Infrastructure

Since 1980s, information and communication technologies (ICT) have been evolving rapidly and now they are key elements in the modern society and economy.

The 5G high speed network will allow up to 10 Gbps of mobile data transmission per device, which is 10 to 100 times faster than current 4G/LTE network. The 5G network will also allow a low latency below 1 ms which is again 10 to 100 times smaller than the current network. This is a significant gain compared to past jumps in speed and delay. Next generation virtual and augmented reality applications pose significant demands in both bandwidth and latency and will be the ones that benefit the most from these features of the 5G network as illustrated in **Fig. 5**. The
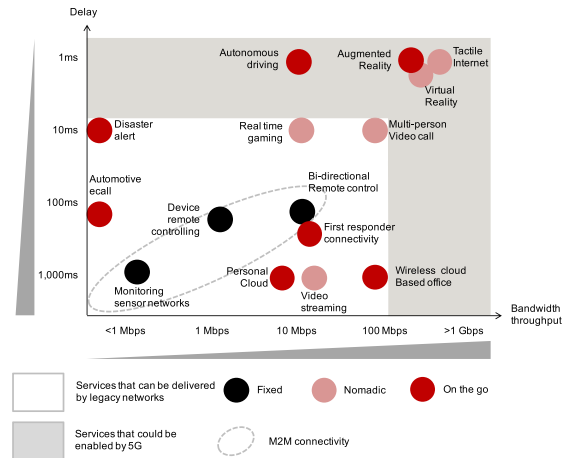

Fig. 5 Bandwidth and latency requirements of potential 5G use cases ((c) GSMA Intelligence 2015 [58]).

number of networked devices will also increase significantly, with the network having a capacity of about 1 Tb/s/km^2. This will allow bandwidth for streaming to tens of thousands of people in a football stadium to broadcast simultaneous, individually tuned HD content.

3.3 Implications for Current Applications

5G infrastructure has significant implications for a number of applications in both the near and far future. Examples include interactive gaming, industrial support, and entertainment.

3.3.1 Gaming and Entertainment

One of the first and perhaps most obvious areas that will benefit from increased network connectivity is that of gaming, especially mobile and location based games. With Pokemon Go showing the power of AR to bring people together on a mass interactive scale, 5G will improve the network dependent mechanisms that support localization of players, interactions between players and the world, and the sharing and joint manipulation of AR content. Low latency haptic devices and controllers that can be used outdoors will open the floodgates for tools, sports equipment, (safe) virtual weapons, and magic abilities that literally are only limited by human imagination. Microtransaction business models will likely be useful for managing the creation and sale of these types of virtual items, much like MMORPGs and other games manage virtual purchases today [19]. This has significant implications for online business as consumers move from tangible goods to digital content and services.

The concept of ePPV, or enhanced pay-per-view [1], may also provide a good business model with which to stream real time virtual experiences with the user. Wireless streaming to multiple devices in theaters can also add personalization and interactivity to current theater experiences, giving everyone a unique perspective on the movie or media being viewed at the time. This will eventually extend to streaming content services in the home and eventually anywhere, so that interactive movies and 3D event reconstructions will be available on the go. Some efforts towards high fidelity reconstructions of images and museums already exist, including the Gigapixel projects [6], [46] and other related media digitization.

3.3.2 Haptic Evolution

Virtual reality technologies are not limited to reproducing visual and auditory sensations. More specifically, haptic interfaces that can reproduce a sense of touch have been intensively investigated for decades. By applying forces and vibrations to the user's sensory organs or by changing the shape and texture of the device surfaces, they can give realistic cutaneous and/or kinesthetic sensations. Haptic interfaces are not only essential for many VR applications, e.g., driving and surgical simulators where a subtle difference of force sensations can result in fatal accidents, but are also desirable in other applications such as sports and remote conferences. One unique aspect of haptic interfaces is that they are mostly bidirectional, i.e., they are both sensory and display devices at the same time, which serves to reproduce and receive forces to and from the user. For example, to present a virtual wall, a haptic device needs to produce a reaction (resistance) force in proportion to the action force exerted by the user. The virtual wall will not feel rigid if the feedback loop is not fast enough. Haptic interfaces can function minimally at slow refresh rates of less than 1,000 Hz and high latencies of 30 to 50 ms, however, extremely fast refresh rates (> 10 KHz) and a low latency (~1 ms) are desirable for high quality, realistic feedback. Aided by the 5G network, future Internet applications will be tactile, where users will be able to touch and manipulate items remotely and exchange nonverbal communication cues like handshaking. This "Haptic Internet" or "Tactile Internet", indicated at the top right of Fig. 5, will add another modality of realism to virtual and augmented reality applications.

3.3.3 Education

Providing a good, effective education is a focus of many, if not all, developed countries. Teachers across the globe strive to provide their students with effective tools for learning and developing in fun and creative ways. Broadcast interfaces such as streaming and archiving of classes on YouTube, massive online open courses (MOOCs) [52], remote education and distance learning programs, and automated / intelligent training programs have recently made leaps and bounds in providing free, consistent education to anyone with an internet connection. In the near future, Virtual reality promises to be able to provide an even more effective way of teaching by integrating rich, detailed content that does not have to be printed on a textbook page. However, just like high speed broadband was an enabling factor in realizing streaming video content, 5G will be an important, if not essential, step in ensuring that the masses have easy, consistent access to virtual educational material, live courses, and telepresent professors and teachers. This new generation of educational interfaces will likely revolutionize the way we think about education in the classroom, and it may be the case that the concept of a joint, physical classroom gradually fades into history. Bringing education to students in this way can help us more effectively distribute skilled teachers to remote locations, eliminate the need for children to travel great distances to school districts, and improve the efficiency of teaching infrastructure.

3.3.4 Healthcare

Healthcare is another major application of virtual and augmented reality systems. With 5G networks in place, a sick person can consult with a family doctor during travel as if they were meeting in the same place. The doctor may immediately share the space with another expert in a completely different office or country and ask for the second opinion in real time with the patient present in the co-located space. If that particular instance required emergency surgery, the physicians could guide the patient to a nearby clinic to perform a remote robot operation. Remote imaging, remote diagnosis and remote surgery are all available today already, however, they are not widespread due to several issues. These include high prices, limited device interoperability, established clinical protocols, and latency associated with bidirectional feedback of surgical tools. 5G can potentially eliminate these barriers, allowing surgeons to perform as if they were on-site, helping propel remote healthcare into the mainstream, and ensuring quality health and medical care on a larger scale than available today.

4. AR/VR in the Era beyond 5G

Realizing the 5G network will have significant implications for many areas of technology and for fields like art and entertainment.

4.1 The Virtual Office

Perhaps one of the most intriguing applications of VR is that of completely integrated telepresence, which has significant implications for office work. With current applications like Skype and GitHub, we have started to see a gradual migration towards work-from-home careers, especially for fields like computing and engineering. The concept of a virtual office in combination with technologies like Fusion 4D [13] are likely to solve many existing problems with telepresence and provide for completely remote office experiences. Moreover, mobile 5G can completely transform how we treat transportation. Being able to stream content and collaborate while riding a train or bus or in the passenger seat of a car can enable employees to conduct business on the go, which in turn can encourage the use of mass and automated transportation methods.

4.2 Reality Remapping

At some point, we will have complete control over the information we present to our eyes using virtual technology. In many cases, such as for individuals that work in drab cubicles, it will be advantageous to have a "change of scene," where the user displays remote mountains instead of walls, and the Milky Way instead of the ceiling. In the case of a portable workspace, we will be able to 'carry' a part of the real office when mobile, or mix parts of the real world from different locations on demand to best cope with the current activities. An individual's surroundings can be tailored by cutting, copying, pasting, scaling, deforming and mashing up different realities. We call this flexible framework 'reality remapping,' and foresee that it will be realized in the era beyond 5G. Though achieving such realistic content is largely a technology rather than network related problem, 5G will enable higher speed sharing of such content. Much like YouTube users share 1080p video content in real time, VR users will share vast 3D worlds in real time. Virtual artists are already coming up with right 3D art using applications like TiltBrush [11]. Sharing of
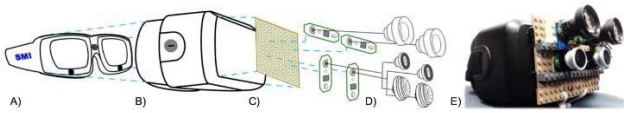
Fig. 6 Image showing the general structure of ModulAR hardware, including the A) an SMI stereoscopic eye-tracker, B) Oculus Rift DK2 head mounted display, C) modular attachment plate used to interchange various camera-lens modules, D) stereo camera pairs with telescopic, fisheye, and ultrawide vision augmentation lenses, and E) an example of a configuration that allows the user to merge a binocular telescopic video stream into a one-to-one field of view [42].

these intricate designs in real-time will enable new collaborative experiences.

4.3   Vision Augmentation

Though AR and VR are typically thought of as technologies that add virtual elements to the real world, another subclass of AR technology is that of vision augmentation (VisAug). This field seeks to improve or modify human vision to provide sight or augmentation that is not possible with the naked eye. The most ancient of these interfaces are spectacles (glasses). Though unrelated to virtual information, glasses embody the essence of VisAug: to provide a better view of the world than with the naked eye. VisAug has gone from from glasses to binoculars, to surgical lenses, and is not being conducted with AR technology. Much like optical elements modify an image, we can now use combinations of optics, algorithms, cameras that utilize different spectral information, and displays to reach new levels of VisAug technology. One such example is the ModulAR display (**Fig. 6**) [42], which provides eye-engaged augmentations with interchangeable lenses. Using a combination of camera-lens modules and rendering techniques, users can engage augmentations with virtual vision augmentation displays [42].

Another example of augmentation that is more software based is that of defocus correction [24]. This algorithm uses a calibrated HMD-camera system to modify background AR information overlaying a particular region of text, thereby providing a clearer rendered image to the user's eyes. Obtaining high resolution scans of the environment via network or cloud services will further enable such technologies to augment or assist human vision. Moreover, having extremely low latency access to 3D cloud data will prevent devices from having to constantly conduct reconstructions or filter vast amounts of augmentative information by facilitating instant access to offloaded processing so that users can have more effective augmentations and digital assistance [9]. In this sense, AR and VisAug devices will become terminals for accessing augmentative information rather than being computing devices. This can also help realize predictive or x-ray vision interfaces. In many cases, traffic or driving in particular, we would like to be able to see additional information about the moving objects around us, other cars, and other incoming vehicles that may be out of sight. Several recent works have looked at automatically predicting trajectory information [25] and determining the best way to utilize cameras to integrate "x-ray vision" information into a user's view [34]. 5G can help integrate the networking, synchronization, and presentation of data, such as sets of data from traffic cameras or surveillance systems, into our visual system in order to see objects or motion well before an accident.

4.4   Beyond Immersive and Hyper-realistic Experiences

Streaming of such high fidelity content begs the question of how far we can take these artificial worlds. Because we can manipulate the user's perception with such content, it is possible that we can come up with something beyond what users typically call "presence," or the feeling and perception that you are present in a remote or virtual space. Once presence is achieved, we can begin designing content that takes advantage of our perceptual deficits in ways that our physical world falls short. Good examples are the various optical illusions we can see on printed paper, like Hering's famous optical illusion (two straight and parallel lines in front of radial background appear as if they were bowed outwards). Though some research has been carried out to date, like using changes in stereo to amplify or emphasize colors or content [61], we still largely have not taken advantage of the capabilities of all that AR and VR have to offer in the domain of changing or augmenting perception.

5.   Conclusion

The 5G network will open a great number of new opportunities for streaming media, telepresence, medicine, and education. Though a number of challenges like latency and bandwidth will still remain in more remote or crowded locations, we will soon have the ability to transmit 3D model or video stream data that will allow users to be completely transported to another world, without wires or tethering. As such, we will likely see a migration to virtual workspaces for education, remote machine operation, telemedicine, and general conferencing applications. Other areas like mobile gaming, vision enhancement, and collaborative art promise to deliver rich environments with which to develop new content and better performance through exciting new human-machine interfaces.

References

[1]   5G Mobile Communications Systems for 2020 and beyond, 5GMF White Paper, *Proc. 5th Generation Mobile Communications Promotion Forum* (2016).

[2]   5G-Reports/5G Service Roadmap 2022. 5F White Paper: 5G Service Roadmap 2022 by the 5G Forum (2016).

[3]   Annett, M., Ng, A., Dietz, P., Bischof, W.F. and Gupta, A.: How low should we go?: Understanding the perception of latency while inking, *Proc. Graphics Interface 2014*, pp.167–174, Canadian Information Processing Society, Chicago (2014).

[4]   Alkhamisi, A.O. and Monowar, M.M.: Rise of augmented reality: Current and future application areas, *International Journal of Internet and Distributed Systems* (2013).

[5]   Alliance, N.G.M.N.: 5G white paper, Next Generation Mobile Networks, Chicago (2015).

[6]   Brady, D.J., Gehm, M.E., Stack, R.A., Marks, D.L., Kittle, D.S., Golish, D.R. and Feller, S.D.: Multiscale gigapixel photography, *Nature*, Vol.486, No.7403, pp.386–389 (2012).

[7]   Billinghurst, M., Weghorst, S. and Furness III, T.: Shared space: An augmented reality approach for computer supported collaborative work, *Virtual Reality*, Vol.3, No.1, pp.25–36 (1998).

[8]   Billinghurst, M. and Kato, H.: Out and About: Real World Teleconferencing, *British Telecom Technical Journal* (*BTTJ*), Millennium Edi-

tion (2000).

[9] Bulling, A., Cakmakci, O., Kunze, K. and Rehg, J.M.: Eyewear Computing–Augmenting the Human with Head-mounted Wearable Assistants, *Dagstuhl Reports*, Vol.6, No.1, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik (2016).

[10] Carlson, C. and Hagsand, O.: DIVE - A Multi-User Virtual Reality System, *Proc. IEEE Virtual Reality Annual International Symposium (VRAIS)*, pp.394–400 (1993).

[11] Cass, S.: Tiltbrush: The killer app for VR, *IEEE Spectrum* (2016).

[12] Caudell, T.P. and Mizell, D.W.: Augmented reality: An application of heads-up display technology to manual manufacturing processes, *Proc. Twenty-Fifth Hawaii International Conference on System Sciences*, Vol.2, pp.659–669, IEEE (1992).

[13] Dou, M., Khamis, S., Degtyarev, Y., Davidson, P., Fanello, S.R., Kowdle, A., Escolano, S.O., Rhemann, C., Kim, D., Taylor, J., Kohli, P., Tankovich, V. and Izadi, S.: Fusion4D: Real-time performance capture of challenging scenes, *ACM Trans. Graphics (TOG)*, Vol.35, No.4, p.114 (2016).

[14] FOV2GO, available from ⟨http://projects.ict.usc.edu/mxr/diy/fov2go/⟩ (accessed 2016-11-07).

[15] Fritsche, L., Unverzag, F., Peters, J. and Calandra, R.: First-person teleoperation of a humanoid robot, *2015 IEEE-RAS 15th International Conference on Humanoid Robots (Humanoids)*, IEEE, pp.997–1002 (2015).

[16] Fettweis, G.P.: A 5G wireless communications vision, *Microwave Journal*, Vol.55, No.12, pp.24–36 (2012).

[17] Furlan, R.: The future of augmented reality: Hololens-Microsoft's AR headset shines despite rough edges [Resources_Tools and Toys], *IEEE Spectrum*, Vol.53, No.6, p.21 (2016).

[18] Greenhalgh, C. and Benford, S.: MASSIVE: A Collaborative Virtual Environment for Teleconferencing, *Transactions on Computer-Human Interaction (TOCHI)*, Vol.2, No.3, pp.239–261, ACM (1995).

[19] Hamari, J. and Lehdonvirta, V.: Game design as marketing: How game mechanics create demand for virtual goods, *International Journal of Business Science & Applied Management*, Vol.5, No.1, pp.14–29 (2010).

[20] Heilig, M.: Sensorama simulator, U.S. Patent 3,050,870 (1962).

[21] Hoggan, E., Brewster, S.A. and Johnston, J.: Investigating the effectiveness of tactile feedback for mobile touchscreens, *Proc. SIGCHI Conference on Human Factors in Computing Systems*, pp.1573–1582, ACM (2008).

[22] Huang, F.C., Chen, K. and Wetzstein, G.: The Light Field Stereoscope: Immersive Computer Graphics via Factored Near-eye Light Field Displays with Focus Cues, *ACM Trans. Graphics (TOG)*, Vol.34, No.4, ACM (2015).

[23] Intelligence, G.S.M.A.: Understanding 5G: Perspectives on future technological advancements in mobile, London, UK (2014).

[24] Itoh, Y. and Klinker, G.: Vision enhancement: Defocus correction via optical see-through head-mounted displays, *Proc. 6th Augmented Human International Conference*, pp.1–8, ACM (2015).

[25] Itoh, Y., Orlosky, J., Kiyokawa, K. and Klinker, G.: Laplacian Vision: Augmenting Motion Prediction via Optical See-Through Head-Mounted Displays, *Proc. 7th Augmented Human International Conference*, ACM (2016).

[26] Leithinger, D., Follmer, S., Olwal, A. and Ishii, H.: Physical telepresence: Shape capture and display for embodied, computer-mediated remote collaboration, *Proc. 27th Annual ACM Symposium on User Interface Software and Technology*, pp.461–470, ACM (2014).

[27] Kanter, T., Fors, U. and Rahmani, R.: Immersive Networking-A Framework for Virtual Environments with Augmented Reality in Human Decision-Making, *International Journal of Multimedia and Ubiquitous Engineering*, Vol.11, No.6, pp.43–60 (2016).

[28] Kato, G., Kuroda, Y., Nisky, I., Kiyokawa, K. and Takemura, H.: HapSticks: A novel method to present vertical forces in tool-mediated interactions by a non-grounded rotation mechanism, *IEEE World Haptics Conference (WHC 2015)*, pp.400–407, IEEE (2015).

[29] Kato, H. and Billinghurst, M.: Marker tracking and hmd calibration for a video-based augmented reality conferencing system, *Proc. 2nd IEEE and ACM International Workshop on Augmented Reality, 1999 (IWAR'99)*, pp.85–94, IEEE (1999).

[30] Keefe, D.F., Feliz, D.A., Moscovich, T., Laidlaw, D.H. and LaViola Jr, J.J.: CavePainting: A fully immersive 3D artistic medium and interactive experience, *Proc. 2001 Symposium on Interactive 3D Graphics*, pp.85–93, ACM (2001).

[31] Kiyokawa, K., Takemura, H. and Yokoya, N.: SeamlessDesign for 3D Object Creation, *IEEE MultiMedia Magazine, ICMCS '99 Special Issue*, Vol.7, No.1, pp.22–33, IEEE (2000).

[32] Kiyokawa, K., Billinghurst, M., Hayes, S.E., Gupta, A., Sannohe, Y. and Kato, H.: Communications Behaviors of Co-located Users in Collaborative AR Interfaces, *Proc. IEEE and ACM International Symposium on Mixed and Augmented Reality 2002 (ISMAR 2002)*, IEEE

(2002).

[33] Leu, M.C., ElMaraghy, H.A., Nee, A.Y., Ong, S.K., Lanzetta, M., Putz, M. and Bernard, A.: CAD model based virtual assembly simulation, planning and training, *CIRP Annals-Manufacturing Technology*, Vol.62, No.2, pp.799–822 (2013).

[34] Maia, L.F., Viana, W. and Trinta, F.: A real-time x-ray mobile application using augmented reality and google street view, *Proc. 22nd ACM Conference on Virtual Reality Software and Technology*, pp.111–119, ACM (2016).

[35] Mekni, M. and Lemieux, A.: Augmented reality: Applications, challenges and future trends, *Applied Computational Science—Proc. 13th International Conference on Applied Computer and Applied Computational Science (ACACOS '14)*, Kuala Lumpur, Malaysia, pp.23–25 (2014).

[36] Milgram, P., Takemura, H., Utsumi, A. and Kishino, F.: Augmented reality: A class of displays on the reality-virtuality continuum, *Photonics for Industrial Applications*, pp.282–292, International Society for Optics and Photonics (1995).

[37] Millar, J.: SIMNET images, available from ⟨http://www.jimmillar.net/BBN.html⟩ (accessed 2016-11-08).

[38] NetWorld2020, E.T.P.: 5G: Challenges, research priorities, and recommendations, Joint White Paper September (2014).

[39] Newcombe, R., Lovegrove, S. and Davison, A.: DTAM: Dense Tracking and Mapping in Real-time, *Proc. IEEE International Conference on Computer Vision (ICCV)* (2011).

[40] Newcombe, R., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A., Kohi, P., Shotton, J., Hodges, S. and Fitzgibbon, A.: KinectFusion: Real-time Dense Surface Mapping and Tracking, *Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (2011).

[41] Ong, S.K. and Nee, A.Y.C.: *Virtual and augmented reality applications in manufacturing*, Springer Science & Business Media (2013).

[42] Orlosky, J., Toyama, T., Kiyokawa, K. and Sonntag, D.: Modular: Eye-controlled vision augmentations for head mounted displays, *IEEE Trans. Visualization and Computer Graphics*, Vol.21, No.11, pp.1259–1268 (2015).

[43] Orts-Escolano, S., Rhemann, C., Fanello, S., Chang, W., Kowdle, A., Degtyarev, Y., Kim, D., Davidson, P.L., Khamis, S., Dou, M., Tankovich, V., Loop, C., Cai, Q., Chou, P.A., Mennicken, S., Valentin, J., Pradeep, V., Wang, S., Kang, S.B., Kohli, P., Lutchyn, Y., Keskin, C. and Izadi, S.: Holoportation: Virtual 3D Teleportation in Real-time, *Proc. 29th Annual Symposium on User Interface Software and Technology*, pp.741–754, ACM (2016).

[44] Palattella, M.R., Dohler, M., Grieco, A., Rizzo, G., Torsner, J., Engel, T. and Ladid, L.: Internet of Things in the 5G Era: Enablers, Architecture, and Business Models, *IEEE Journal on Selected Areas in Communications*, Vol.34, No.3, pp.510–527 (2016).

[45] Petersen, N., Pagani, A. and Stricker, D.: Real-time modeling and tracking manual workflows from first-person vision, *Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp.117–124, IEEE (2013).

[46] Proctor, N.: The Google Art Project: A New Generation of Museums on the Web?, *Curator: The Museum Journal*, Vol.54, No.2, pp.215–221 (2011).

[47] Schmalstieg, D., Fuhrmann, A., Hesina, G., Szalavári, Z., Encarnaçao, L.M., Gervautz, M. and Purgathofer, W.: The studierstube augmented reality project, *Presence: Teleoperators and Virtual Environments*, Vol.11, No.1, pp.33–54 (2002).

[48] Shapira, L. and Freedman, D.: Reality Skins: Creating Immersive and Tactile Virtual Environments, *Proc. International Symposium on Mixed and Augmented Reality* (2016).

[49] Soldani, D., Virette, D., Cordara, G. and Zhang, Q.: An Enriched Multimedia Experience for Wireless Networks in Horizon 2020 and Beyond, *IEEE COMSOC MMTC E-Letter*, Vol.7, No.8 (2012).

[50] Sweeney, C., Höllerer, T. and Turk, M.: Improved outdoor augmented reality through "Globalization", *Proc. IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp.1–4, IEEE (2013).

[51] Sutherland, I.E.: A head-mounted three dimensional display, *Proc. Fall Joint Computer Conference, Part I*, pp.757–764, ACM (1968).

[52] Stathakarou, N., Zary, N. and Kononowicz, A.A.: Beyond xMOOCs in healthcare education: Study of the feasibility in integrating virtual patient systems and MOOC platforms, *PeerJ*, Vol.2, e672 (2014).

[53] Takala, T.M. and Matveinen, M.: Full body interaction in virtual reality with affordable hardware, *Proc. IEEE Virtual Reality (VR)*, pp.157–157, IEEE (2014).

[54] Thomas, B., Close, B., Donoghue, J., Squires, J., De Bondi, P., Morris, M. and Piekarski, W.: ARQuake: An outdoor/indoor augmented reality first person application, *Proc. 4th International Symposium on Wearable computers*, pp.139–146, IEEE (2000).

[55] Thorpe, J.A., Bloedorn, G.W., Taylor, R., Miller, D.C. and Cyrus, M.: *The SIMNET network and protocol*, BBN LABS INC CAMBRIDGE

MA (1987).

[56] Vate-U-Lan, P.: Envisioning Augmented Reality: Smart Technology for the Future, *Proc. 8th International Conference on eLearning for Knowledge-Based Society* (2012).

[57] Wang, X., Kim, M.J., Love, P.E. and Kang, S.C.: Augmented Reality in built environment: Classification and implications for future research, *Automation in Construction*, Vol.32, pp.1–13 (2013).

[58] Warren, D. and Dewar, C.: *Understanding 5G: Perspectives on future technological advancements in mobile*, London: GSMA Intelligence (2014).

[59] Weinbaum, S.G.: *Pygmalion's Spectacles*, Booklassic (2015).

[60] Westerfield, G., Mitrovic, A. and Billinghurst, M.: Intelligent augmented reality training for motherboard assembly, *International Journal of Artificial Intelligence in Education*, Vol.25, No.1, pp.157–172 (2015).

[61] Zhang, H., Cao, X. and Zhao, S.: Beyond stereo: An exploration of unconventional binocular presentation for novel visual experience, *Proc. SIGCHI Conference on Human Factors in Computing Systems*, pp.2523–2526, ACM (2012).

[62] Zyda, M.J., McGhee, R.B., McConkle, C.M., Nelson, A.H. and Ross, R.S.: A real-time, three-dimensional moving platform visualization tool, *Computers & Graphics*, Vol.14, No.2, pp.321–333 (1990).

Haruo Takemura received B.E., M.E., and Ph.D. degrees from Osaka University in 1982, 1984, and 1987, respectively. In 1987, he joined Advanced Telecommunication Research Institute, International. In 1994, he joined Nara Institute of Science and Technology, as an associate professor in the Graduate School of Information Science and Technology. From 1998 to 1999, he was a visiting associate professor at the University of Toronto, Ontario, Canada. He is a professor at Cybermedia Center, Osaka University since 2001. His research interests include interactive computer graphics, human-computer interaction, and mixed reality. He is a member of IEICE, IPSJ, VRSJ, the HIS, IEEE, and ACM.

Jason Orlosky received a bachelor's degree in computer engineering from the Georgia Institute of Technology in 2006. He then worked in information technology for 3 years in the United States before returning to school in 2010 to study Japanese language and literature. In 2011, he studied abroad as a research student at Osaka University, where he then became a Ph.D. student in 2013. After receiving his Ph.D. in 2016, he received a fellowship from the Japan Society for the Promotion of Science from 2015 – 2017, and is currently a postdoctoral researcher at Osaka University studying Augmented Reality and Artificial Intelligence.

Kiyoshi Kiyokawa received his M.E. and Ph.D. degrees in information systems from Nara Institute of Science and Technology in 1996 and 1998, respectively. He is currently an associate professor at Cybermedia Center, Osaka University. He worked for Communications Research Laboratory from 1999 to 2002. He was a visiting researcher at Human Interface Technology Laboratory at the University of Washington from 2001 to 2002. His research interests include virtual reality, augmented reality, and 3D user interfaces.