



# Grand Challenges for Augmented Reality

Mark Billinghurst<sup>1,2\*</sup>

<sup>1</sup>STEM, University of South Australia, Mawson Lakes, SA, Australia, <sup>2</sup>Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand

**Keywords:** augmented reality, grand challenge, display, interaction, tracking, collaboration, ethics

## INTRODUCTION

In his 1965 article, *The Ultimate Display*, Ivan Sutherland imagined a future computer interface that blurred the separation between the digital and physical worlds (Sutherland, 1965). At the time, he was making this vision a reality, creating a see-through head mounted display (HMD) that allowed users to see virtual images superimposed over the real world (Sutherland, 1968). The user's head position was tracked, so the virtual content appeared fixed in space, and a handheld wand could be used to interact with it.

Although the term was not coined until decades later, Sutherland's system was the first working Augmented Reality (AR) interface. AR is technology with three key characteristics (Azuma, 1997); 1) it combines real and virtual images, 2) is interactive in real time, and 3) the virtual imagery is registered in three dimensions. Sutherland's work had these properties, but over 50 years later, his vision of the *Ultimate Display* still hasn't been achieved and more research is needed.

Azuma's definition of AR provides guidance on the technology required to create an AR experience. In order to combine real and virtual images display technology is needed. To support interaction in real time user interface technologies are required. To register AR content in three dimensions tracking technology is needed.

Once these technologies were only available in research labs, but today they are available in people's hands. Current mobile phones with cameras, GPS and inertial sensors, high resolution screens, fast networking and powerful CPUs and graphics processors are the most common way that people experience AR. Compatible with hundreds of millions of devices, Apple's ARKit (Apple, 2020), and Google's ARCore (Google, 2020a) provide accurate AR tracking for mobiles. A user can look at the camera view on their phone screen and see virtual objects in their real world. Mobile AR applications such as *Pokemon Go* have been downloaded over a billion times (NintendoSoup, 2019), showing how readily accessible the technology is.

However, the user experience provided by a phone is very different from the Sutherland's vision of hands-free interaction, stereo graphics, and virtual imagery always in a person's field of view. Mobile AR provides an easily accessible entry point, but the true potential of AR is achieved through using head mounted displays, with richer interaction and better tracking techniques. In each of these areas there are important Grand Challenges that need research, as discussed below.

## Research in Display Technology

Sutherland used miniature cathode ray tubes mounted on the head with optical combiners to create a stereo see-through AR display. However, this had a limited field of view, resolution and refresh rate. One Grand Challenge is to create a wide field of view, high resolution, see-through display in a socially acceptable form factor. There are a number of factors that need to be addressed before HMDs can become a replacement for smartphones. These include creating a sunglass like form factor, providing sufficient brightness and contrast, having a high resolution and wide field of view, addressing eyestrain, and enabling people to see each other's eyes (Azuma, 2017). Research is

## OPEN ACCESS

### Edited and reviewed by:

Mel Slater,  
University of Barcelona, Spain

### \*Correspondence:

Mark Billinghurst  
mark.billinghurst@unisa.edu.au

### Specialty section:

This article was submitted to  
Augmented Reality,  
a section of the journal  
Frontiers in Virtual Reality

**Received:** 30 June 2020

**Accepted:** 28 January 2021

**Published:** 05 March 2021

### Citation:

Billinghurst M (2021) Grand Challenges  
for Augmented Reality.  
*Front. Virtual Real.* 2:578080.  
doi: 10.3389/frvir.2021.578080

ongoing in many of these areas. For example, a pinhole screen can be used to create a wide field of view see-through AR display (Maimone et al., 2014) and holographic projection can be used to achieve full color, high contrast AR images in an eye-glass form factor (Maimone et al., 2017).

Other areas are also important, such as the vergence accommodation problem caused by a display only having a single focal plane, preventing people from keeping the AR content in focus while also focusing on objects in the real world at a different distance. Variable focal planes can enable users to view virtual content at different focal lengths (Liu et al., 2008). Light Field Displays and light fields provide one way to show photorealistic content to the user and are a prerequisite for creating “True Augmented Reality” (Sandor et al., 2015). There are also interesting innovations happening in the commercial sector, such as from companies like Mojo Vision (Mojo Vision, 2020) who are developing AR enabled contact lenses, but these are many years away from commercialization.

## Research in Interaction

Sutherland’s system supported simple interaction with a handheld wand. Another Grand Challenge is to enable people to interact with AR content as easily as they do with real objects. Many researchers are exploring natural user interfaces such as using tangible objects to interact with AR content (Tangible AR interfaces (Billinghurst et al., 2008)) or free-hand gesture manipulation (Sharp et al., 2015). Modern AR displays such as the HoloLens2 (Microsoft, 2020a) support natural two-handed gesture input, allowing people to reach out and grab virtual content. However, it is possible to go beyond this and combine speech and gesture together to create multimodal interfaces where the strengths of one modality compensates for the weakness of another (Nizam et al., 2018). Addition of eye-tracking, full-body input, and other non-verbal cues can provide even more intuitive multimodal interaction. Research also needs to be conducted into interaction methods using techniques not possible in the real world. Brain computer interaction methods enable brain activity to select AR content (Si-Mohammad et al., 2018), and other physiological sensors can enable AR to respond to user heart rate or emotional state. There are many opportunities to create even better AR interaction methods.

## Research in Tracking

A key feature of AR systems is that the content appears to be fixed in space, which requires the user’s viewpoint to be continuously tracked. Sutherland achieved this by using mechanical and ultrasonic trackers to measure where the user’s HMD was and render the virtual imagery from that same position. Tracking technology has improved significantly, but another Grand Challenge is to precisely locate a user’s position in any location. There has been a significant amount of research on computer vision methods for tracking user viewpoint without knowing any visual features (Kim et al., 2018). Hybrid approaches that combine vision-based SLAM tracking with GPS and inertial sensors can be used for a more robust result (Liu et al., 2016). However, one area that hasn’t been well explored are hybrid

approaches for very large-scale tracking. Wide area tracking can be achieved using sensor fusion from a dynamic combination of mobile and stationary tracking (Pustka and Klinker, 2008). Deep Learning could be used to coordinate multiple tracking systems and provide some scene understanding (Garon and Lalonde, 2017). Finally, there is a recent trend toward AR cloud-based tracking where features captured by a user’s device are uploaded to the cloud and fused to provide a ubiquitous tracking service. HoloRoyale is one of the first examples of using city scale AR tracking from an AR cloud service to enable collaborative gaming (Rompapas et al., 2019). Commercial software from companies such as Ubiquity6 (Ubiquity6, 2020) enable large scale AR cloud tracking. However, none of these systems yet provide large-scale precise tracking, so more work is needed.

## Research in Perception and Neuroscience

In addition to Grand Challenges in fundamental technology, there are other areas of AR that need to be addressed, such as exploring perceptual and neuroscience issues. AR systems create an illusion to convince the brain that virtual content actually exists in the real world. There are a number of perceptual problems that can occur in AR, classified into environmental, capturing, augmentation, display device, and user issues (Kruijff et al., 2010). Considerable research has been conducted on how to make AR content appear the same as real objects, including the use of virtual lighting (Agusanto et al., 2003), shadows (Sugano et al., 2003), real object occlusion (Breen et al., 1996) and similar methods. The goal is to create digital objects that have strong “Object Presence” and appear to be really there (Stevens and Jerrams-Smith, 2000). However, unlike Presence in Virtual Reality, Object Presence in AR has not been well studied. Most of these systems are evaluated using subjective measures, but EEG can be used as an objective measure to evaluate the quality of experience (Bauman and Seeling, 2018). EEG could also be used to explore the cognitive load of using AR interfaces, measure emotional response to AR stimuli, monitor shared brain activity in collaborative AR experiences, and more. So, there is significant opportunity to use neuroscience to understand the perceptual and psychological basis of AR.

## Research in Collaboration

There are also many application areas that could be studied in more detail. One important area is using AR to enable remote people to work together as easily as if they were face to face. Early experiments showed that AR views of video avatars provided a significantly higher degree of Social Presence than traditional video conferencing (Billinghurst and Kato, 2002). More recently, Microsoft’s Holoportation captured full 3D models of people in real time and showed them as life-sized AR avatars in a user’s real environment, enabling the sharing of rich communication cues (Orts-Escolano et al., 2016). The company Spatial provides a commercial application that can superimpose AR avatars over the real world in a very natural way (Spatial.io, 2020).

There are also many examples of wearable AR systems can be used to enable a remote expert to see through a local user’s eyes and provide AR cues to help them perform real-world tasks (Kim et al., 2019). Microsoft’s Remote Assist product (Microsoft, 2020b), and others, have made this type of experience commercially available. The emerging field of Empathic

Computing (Piumsomboon et al., 2017) goes beyond this to explore how physiological cues can be combined with AR in collaborative interfaces to enable remote people to share what they are seeing, hearing and feeling. There is also opportunity to study how to support viewing large scale social networks in AR interfaces, including using visual and spatial cues to separate out dozens of social contacts (Nassani et al., 2017). However, there is still very little research conducted on collaborative AR. A survey of 10 years of user studies until 2015, found that only 15 of the 369 AR studies reviewed were collaborative studies, and only seven of these used AR HMDs (Dey et al., 2018).

## Research in Social and Ethical Issues

Finally, there are social and ethical issues that need to be addressed. The difficulty of Google Glass (Google, 2020b) and other AR displays to get consumer acceptance, shows that widespread use of HMD-based AR may depend more on social than technical issues. Rauschnabel explored the technology acceptance drivers of AR smart glasses (Rauschnabel, and Ro, 2016), while Pascoal studied acceptance in outdoor environments (Pascoal et al., 2018).

When AR devices become more widely used a number of ethical issues may arise. Who should be allowed to place AR content in the view of a person and what are the ethics around AR advertising? What is the consequence of people having different views of the same real environment? Brinkman discusses the privacy implications of AR as an extension of the home and AR advertising (Brinkman, 2014). Pase lists a number of questionable

ethical uses of pervasive AR, such as deception, surveillance, behavior modification, and punishment (Pase et al., 2012). AR technology could be used to create mediated reality experiences, removing from view certain parts of the real world, which could have public safety issues (Mann, 2002). Users capturing and sharing their surroundings for AR cloud tracking or remote collaboration could also raise significant concerns. Wasson has written about the legal, ethical and privacy issues of AR (Wasson, 2014), but there is still much more research needed.

## CONCLUSION

Over 50 years ago Sutherland provided a compelling vision of how the physical and digital worlds could be seamlessly combined together. However, there is still significant research that needs to be done to make this vision a reality. Grand Challenges exist in fundamental display, interaction and tracking technologies, and also the perception/neuroscience of AR, using AR for collaboration, and exploring the social and ethical aspects. Addressing these topics will enable Augmented Reality to reach its full potential as a transformative technology.

## AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

## REFERENCES

- Agusanto, K., Li, L., Chuangui, Z., and Sing, N. W. (2003). "Photorealistic rendering for augmented reality using environment illumination," in The second IEEE and ACM international symposium on mixed and augmented reality, Tokyo, Japan, October 10, 2003 (Piscataway, New Jersey, USA: IEEE), 208–216.
- Apple (2020). *ARKit*. Available at: <https://developer.apple.com/augmented-reality/> (Accessed June 27, 2020).
- Azuma, R. (1997). A survey of augmented reality. *Presence* 6 (4), 355–385. doi:10.1162/pres.1997.6.4.355
- Azuma, R. (2017). "Making augmented reality a reality," in Applied industrial optics: spectroscopy, imaging and metrology, San Francisco, CA, June 26–29, 2019 (Washington, D.C, USA: Optical Society of America), JTu1F-1.
- Bauman, B., and Seeling, P. (2018). "Evaluation of EEG-based predictions of image QoE in augmented reality scenarios," in 2018 IEEE 88th vehicular technology conference (VTC-Fall), Chicago, IL, August 27–30, 2018 (Piscataway, New Jersey, USA: IEEE), 1–5.
- Billinghurst, M., and Kato, H. (2002). Collaborative augmented reality. *Commun. ACM* 45 (7), 64–70. doi:10.1145/514236.514265
- Billinghurst, M., Kato, H., and Poupyrev, I. (2008). Tangible augmented reality. *ACM SIGGRAPH Asia* 7 (2), 1–10. doi:10.1145/1508044.1508051
- Breen, D., Whitaker, R., Rose, E., and Tuceryan, M. (1996). Interactive occlusion and automatic object placement for augmented reality. *Comput. Graphics Forum* 15 (3), 11–22. doi:10.1111/1467-8659.1530011
- Brinkman, B. (2014). "Ethics and pervasive augmented reality: some challenges and approaches," in *Emerging pervasive information and communication technologies (PICT). Law, governance and technology series*. Editor K. Pimple (Dordrecht: Springer), Vol. 11, 149–175.
- Dey, A., Billinghurst, M., Lindeman, R., and Swan, J. (2018). A systematic review of 10 years of augmented reality usability studies: 2005 to 2014. *Front. Robot AI* 5, 37. doi:10.3389/frobt.2018.00037
- Garon, M., and Lalonde, J. (2017). Deep 6-DOF tracking. *IEEE Trans. Vis. Comput. Graph* 23 (11), 2410–2418. doi:10.1109/TVCG.2017.2734599
- Google (2020a). ARCore. Available at: <https://developers.google.com/ar> (Accessed December 15, 2020).
- Google (2020b). Google glass. Available at: <https://www.google.com/glass/start/> (Accessed February 4, 2020).
- Kim, K., Billinghurst, M., Bruder, G., Duh, H., and Welch, G. (2018). Revisiting trends in augmented reality research: a review of the 2nd decade of ISMAR (2008–2017). *IEEE Trans. Vis. Comput. Graph* 24 (11), 2947–2962. doi:10.1109/TVCG.2018.2868591
- Kim, S., Lee, G., Huang, W., Kim, H., Woo, W., and Billinghurst, M. (2019). "Evaluating the combination of visual communication cues for HMD-based mixed reality remote collaboration," in Proceedings of the 2019 CHI conference on human factors in computing systems, Glasgow Scotland, United Kingdom, May, 2019 (New York, NY: Association for Computing Machinery), 1–13.
- Kruijff, E., Swan, J. E., and Feiner, S. (2010). "Perceptual issues in augmented reality revisited," in IEEE international symposium on mixed and augmented reality, Seoul, Korea, October 13–16, 2010 (Piscataway, New Jersey, USA: IEEE), 3–12.
- Liu, H., Zhang, G., and Bao, H. (2016). "Robust keyframe-based monocular SLAM for augmented reality," in IEEE international symposium on mixed and augmented reality (ISMAR), Merida, Mexico, September 19–23, 2016 (IEEE), 1–10.
- Liu, S., Cheng, D., and Hua, H. (2008). "An optical see-through head mounted display with addressable focal planes," in 2008 7th IEEE/ACM international symposium on mixed and augmented reality, Cambridge, United Kingdom, September 15–19, 2008 (IEEE), 33–42.
- Maimone, A., Georgiou, A., and Kollin, J. (2017). Holographic near-eye displays for virtual and augmented reality. *ACM Trans. Graph.* 36 (4), 1–16. doi:10.1145/3072959.3073624
- Maimone, A., Lanman, D., Rathinavel, K., Keller, K., Luebke, D., and Fuchs, H. (2014). "Pinlight displays: wide field of view augmented reality eyeglasses using defocused point light sources," in ACM SIGGRAPH 2014 emerging

- technologies, Vancouver, Canada, August, 2014 (New York, NY: Association for Computing Machinery), 1.
- Mann, S. (2002). *Mediated reality with implementations for everyday life*. Presence Connect, 1.
- Microsoft (2020a). *Hololens2*. Available at: <https://www.microsoft.com/en-us/hololens/> (Accessed December 7, 2019).
- Microsoft (2020b). *Remote assist*. Available at: <https://dynamics.microsoft.com/mixed-reality/remote-assist/> (Accessed May 28, 2020).
- Mojo Vision (2020). *Moja vision*. Available at: <https://www.mojo.vision/> (Accessed April 29, 2020).
- Nassani, A., Lee, G., Billinghurst, M., Langlotz, T., and Lindeman, R. (2017). "Using visual and spatial cues to represent social contacts in AR," in SIGGRAPH asia 2017 mobile graphics and interactive applications, Bangkok, Thailand, November 2017 (New York, NY: Association for Computing Machinery), 1–6.
- NintendoSoup (2019). *Pokemon go officially hits 1 billion downloads worldwide*. Available at: <https://nintendosoup.com/pokemon-go-officially-hits-1-billion-downloads-worldwide/> (Accessed April 11, 2019).
- Nizam, S., Abidin, R., Hashim, N., Lam, M., Arshad, H., and Majid, N. (2018). A review of multimodal interaction technique in augmented reality environment. *Int. J. Adv. Sci. Eng. Inf. Technol.* 8 (4–2), 1460. doi:10.18517/ijaseit.8.4-2.6824
- Orts-Escolano, S., Rhemann, C., Fanello, S., Chang, W., Kowdle, A., Degtyarev, Y., and Tankovich, V. (2016). "Holoportation: virtual 3D teleportation in real-time," in Proceedings of the 29th annual symposium on user interface software and technology, Tokyo, Japan, October 2016 (New York, NY: Association for Computing Machinery), 741–754.
- Pascoal, R., Alturas, B., de Almeida, A., and Sofia, R. (2018). "A survey of augmented reality: making technology acceptable in outdoor environments," in 2018 13th Iberian conference on information systems and technologies. CISTI, Caceres, June 13–16, 2018 (Piscataway, New Jersey, USA: IEEE), 1–6.
- Pase, S. (2012). "Ethical considerations in augmented reality applications," in Proceedings of the international conference on e-learning-business, enterprise information systems, and e-Government EEE. (The Steering Committee of the World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp)), 1. Las Vegas, Nevada, USA, July 16th - 19th, 2012.
- Piumsomboon, T., Lee, Y., Lee, G., Dey, A., and Billinghurst, M. (2017). "Empathic mixed reality: sharing what you feel and interacting with what you see," in International symposium on ubiquitous virtual reality (ISUVR), Nara, June 27–29, 2017 (Piscataway, New Jersey, USA: IEEE), 38–41.
- Pustka, D., and Klinker, G. (2008). "Dynamic gyroscope fusion in ubiquitous tracking environments," in 2008 7th IEEE/ACM international symposium on mixed and augmented reality, Cambridge, United Kingdom, September 15–18, 2008 (Piscataway, New Jersey, USA: IEEE), 13–20.
- Rauschnabel, P., and Ro, Y. (2016). Augmented reality smart glasses: an investigation of technology acceptance drivers. *Int. J. Technol. Mark.* 11 (2), 123–148. doi:10.1504/IJTMKT.2016.075690
- Rompapas, D. C., Sandor, C., Plopski, A., Saakes, D., Shin, J., Taketomi, T., et al. (2019). Towards large scale high fidelity collaborative augmented reality. *Comput. Graph.* 84, 24–41. doi:10.1016/j.cag.2019.08.007
- Sandor, C., Fuchs, M., Cassinelli, A., Li, H., Newcombe, R., Yamamoto, G., et al. (2015). Breaking the barriers to true augmented reality. *arXiv:1512.05471*.
- Sharp, T., Keskin, C., Robertson, D., Taylor, J., Shotton, J., Kim, D., et al. (2015). "Accurate, robust, and flexible real-time hand tracking," in Proceedings of the 33rd annual ACM conference on human factors in computing systems, South Korea, April 2015 (New York, NY: Association for Computing Machinery), 3633–3642.
- Si-Mohammed, H., Petit, J., Jeunet, C., Argelaguet, F., Spindler, F., Evain, A., et al. (2018). Towards BCI-based interfaces for augmented reality: feasibility, design and evaluation. *IEEE Trans. Vis. Comput. Graph.* 26 (3), 1608–1621. doi:10.1109/TVCG.2018.2873737
- Spatial.io (2020). *Spatial*. Available at: <https://spatial.io/> (Accessed May 1, 2020).
- Stevens, B., and Jerrams-Smith, J. (2000). "The sense of object-presence with projection-augmented models," in International workshop on haptic human-computer interaction, Glasgow, United Kingdom, August–September 31–1, 2000 (Berlin, Heidelberg: Springer), 194–198.
- Sugano, N., Kato, H., and Tachibana, K. (2003). "The effects of shadow representation of virtual objects in augmented reality," in The second IEEE and ACM international symposium on mixed and augmented reality, Tokyo, Japan, October 10, 2003 (Piscataway, New Jersey, USA: IEEE), 76–83.
- Sutherland, I. (1968). "A head-mounted three dimensional display," in Fall joint computer conference (Fall, part I), San Francisco, California, December 9–11, 1968 (New York, NY: ACM Press), 757–764.
- Sutherland, I. (1965). The ultimate display. *Proc. IFIP Congress 2*, 506–508.
- Ubiquity6 (2020). *Ubiquity6*. Available at: <https://ubiquity6.com/> (Accessed January 14, 2020).
- Wassom, B. (2014). *Augmented reality law, privacy, and ethics: law, society, and emerging AR technologies*. Waltham, Massachusetts, USA: Syngress, 360.

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

Copyright © 2021 Billinghurst. 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.