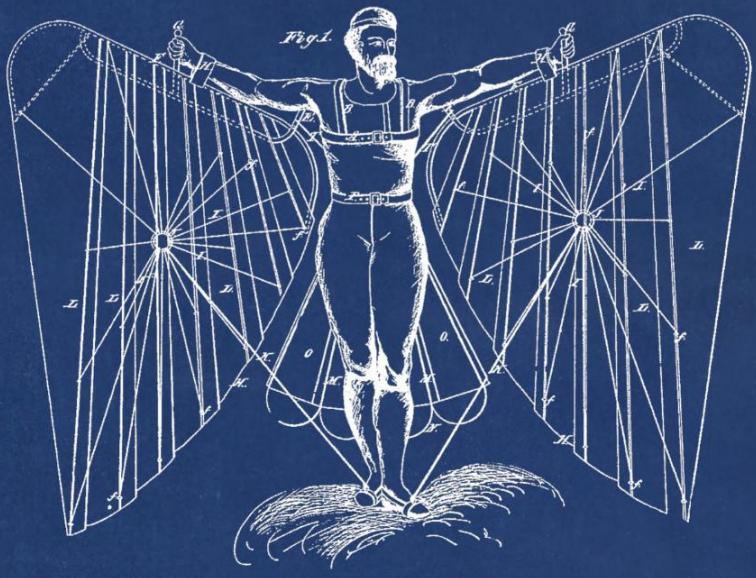
O'REILLY®

AUGMENTED HUMAN

HOW TECHNOLOGY IS SHAPING THE NEW REALITY



HELEN PAPAGIANNIS

Augmented Human

How Technology Is Shaping the New Reality

Helen Papagiannis

Beijing • Boston • Farnham • Sebastopol • Tokyo O'REILLY®

Augmented Human
by Helen Papagiannis

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (*http://oreilly.com/safari*). For more information, contact our corporate/institutional sales department: 800-998-9938 or *corporate@oreilly.com*.

Editors: Susan Conant and Jeff Bleiel
Production Editor: Nicholas Adams
Copyeditor: Bob Russell, Octal Publishing Inc.
Proofreader: James Fraleigh
Indexer: Wendy Catalano
Interior Designer: Monica Kamsvaag
Cover Designer: Edie Freedman
Illustrator: Rebecca Demarest

August 2017: First Edition

Revision History for the First Edition
2017-08-10: First Release

See *http://oreilly.com/catalog/errata.csp?isbn=9781491928325* for release details.

978-1-491-92832-5

[LSI]

Contents

Foreword

When Helen told me that she was authoring *Augmented Human: How Technology is Shaping the New Reality*, I volunteered (actually begged her) to let me write this foreword. At the time, I had not read a manuscript and had only an inkling of what she was going to say, but knowing Helen's scholarly reputation and thoughtfulness, I knew that her take on the emerging technology of augmented reality and its applications in augmenting humans would be insightful, if not profound.

There have been a lot of people who have worked together to create this new medium. For many years it was looked upon as a novelty or solution looking for a problem to solve. There wasn't enough crowd appeal to move it past the tipping point of mass adoption—this being necessary for enough investment and attention toward making AR economically viable, and thus, widespread. My qualification for writing this foreword stems from the dubious honor of being one of those so-called pioneers who suffered through the beginnings of AR technology development.

In my case the journey began 52 years ago, when as a USAF officer at Wright-Patterson Air Force Base, I worked on designing better cockpits for fighter aircraft and other military vehicles. My job was to solve the problem of getting bandwidth to and from the pilot's brain given the sheer complexity of the systems they needed to operate in highly stressful and dangerous environments. This problem is what motivated me to explore augmented reality approaches that might increase pilot awareness of the relationship of their aircraft and to the real world. The idea was to organize and portray information in the form of virtual images projected and superimposed over the real world via helmet worn devices. Later I extended this work to include virtual reality.

Now that AR and VR technology is finally maturing (it took longer than I had thought) my role as a tool builder is coming to an end. It is time to hand the torch to Helen and her cohort to make it do something useful. It is one thing to build a new medium but another thing to put a message in that medium. In the end, the message is more important.

As I anticipated, Dr. Papagiannis has done a splendid job in this authoritative book. It is short but packs a wallop. She has laid a foundation for us by describing what we mean by augmented reality with its various modalities, developed a taxonomy to help us categorize applications and introduced us to the content and applications (along with their pioneers). But she has gone a lot further. In connecting the dots, Helen leads us to realize that AR is not going to be business as usual. This is not about a new medium—it is about augmenting humans. It is not about separating us from experiencing the real world, as television, film and even virtual reality require, but instead, allows technology to blend with the real world, enhancing us in that total blended experience. From this perspective she shows us that AR has the power to empower, and that we need to expand our thinking (and imagination) of what this augmentation can mean to being human in the future.

When it comes to the message delivered by the AR medium, there will be many new degrees of freedom such as non-linear storytelling and alternative realities that change even the physics of experience. As Helen explains we need to throw out the old rulebook of how we understand and interact within the real world. As we embrace these new experiences and their saliencies she envisions an unleashing of discovery, creativity, imagination and what it may mean to be an augmented human in the days to come. No doubt hers is not the final word, but it is an essential and necessary touchstone as we embark on this human augmentation odyssey.

I especially appreciate Helen's insight and sensitivity about the role of artists (or "wonderment operators" as she calls them) as a stimulus to emergence. My own experience is that there is not one community that owns this space. It is as much a place for storytellers and artists as it is for engineers and computer scientists. The experience is what we will remember and will change us. Hopefully the technology (my contribution) becomes "invisible" and doesn't get in the way.

I take to heart Helen's final charge at the conclusion of this book that we need to work together as a civilization to use the tools of our age to lift humanity and inspire a positive change in the world. For in the end, we must answer the question: does augmenting us make our lives better?

—Tom Furness,
Grandfather of AR/VR and Founder
of the Virtual World Society, Seattle,
July 16, 2017

Preface

Why I Wrote This Book

Twelve years ago, I caught my first glimpse of the power of Augmented Reality (AR) as a new communication medium. It was pure magic: a virtual 3-D cube appeared in my physical surroundings and I was awestruck. The augmented cube demo wasn't interactive at the time (it did nothing else other than appear), however, it ignited my imagination for how AR could grow and evolve. At that moment, I dedicated my creative work, research, and public speaking to the new experiences AR made possible.

I wrote this book because I began to witness a much-needed shift from a focus on the technology alone to a push towards creating compelling content and meaningful experiences in AR. This book is about exploring those big ideas and the extraordinary new reality AR affords. Now is the time to dream, design, and build our wondrous future.

As AR advances, we must ask: How can we design AR experiences to enhance a user's life and make it easier and better? MIT Media Lab founder Nicholas Negroponte said, "Computing is not about computers anymore. It is about living." AR is no longer just about the technology, it's about living in the real world, and creating magical and meaningful experiences that are human-centred. This book is about how AR will enrich our daily lives and extend humanity in unprecedented ways.

Who Should Read This Book

It's not too often an entire new medium emerges. You should read this book if you're a maker, a doer, and an explorer who is excited by creating a path where there is no trail, and want to contribute to this rapidly growing industry. You

should also read this book as an informed consumer for a peek at the new experiences that will change the way we live, work, and play.

You are a designer, a developer, an entrepreneur, a student, an educator, a business leader, an artist, and a technology enthusiast curious about and excited by the possibilities AR presents. You are committed to designing and supporting AR experiences for the deepest of human values to have a profound impact on bettering humanity.

No prior knowledge of AR is required to read this book. To get the most out of this book, I do recommend trying out an AR experience first-hand (several ideally), including any of the examples referenced throughout the chapters.

Navigating This Book

This book is organized as follows:

Chapter 1 revisits the classic definition of AR from 1997, expanding on how AR is changing today and beyond. This chapter introduces the next wave of AR that enables a new spatial understanding and sensory awareness to create a more immersive, integrated, and interactive experience.

Chapter 2 explores how computer vision is giving us new eyes and perspectives to engage with the world, from artist installations, to robots and self-driving cars, to assisting people who are vision-impaired.

Chapter 3 investigates research and innovation in haptic technology (touch feedback) to sync what we see with how something feels, as well as creating new ways to communicate using tactility.

In addition to audio used for navigation and narration, Chapter 4 explores approaches to augmented audio and "Hearables" (wearable technology worn in the ear) that make it possible to alter the way you hear your environment, and how your environment even "listens" to you.

In Chapter 5 we learn how digital smell and taste is a growing area of research, prototypes, and product design that can augment the way we share and receive information, enhance entertainment experiences, deepen our understanding of a place, and affect our overall wellbeing.

Chapter 6 looks at how AR is moving past novelty to create compelling storytelling experiences, noting where we've been with recurring storytelling themes and conventions, and where we're headed with emerging styles and mechanisms.

Chapter 7 queries how avatars, intelligent agents, objects, and materials are becoming living contextual change agents: learning, growing, predicting, and shape-shifting to your context.

Chapter 8 surveys the way we are augmenting our bodies from electronic textiles, to technology embedded in the body, to brain-controlled interfaces.

Chapter 9 identifies ten AR experiential categories to date, with the intention to grow the possibilities into the near-future and beyond with a sense of wonder and a commitment to uplifting humanity.

O'Reilly Safari

Safari (formerly Safari Books Online) is a membership-based training and reference platform for enterprise, government, educators, and individuals.

Members have access to thousands of books, training videos, Learning Paths, interactive tutorials, and curated playlists from over 250 publishers, including O'Reilly Media, Harvard Business Review, Prentice Hall Professional, Addison-Wesley Professional, Microsoft Press, Sams, Que, Peachpit Press, Adobe, Focal Press, Cisco Press, John Wiley & Sons, Syngress, Morgan Kaufmann, IBM Redbooks, Packt, Adobe Press, FT Press, Apress, Manning, New Riders, McGraw-Hill, Jones & Bartlett, and Course Technology, among others.

For more information, please visit *http://oreilly.com/safari*.

How to Contact Us

Please address comments and questions concerning this book to the publisher:

O'Reilly Media, Inc.
1005 Gravenstein Highway North
Sebastopol, CA 95472
800-998-9938 (in the United States or Canada)
707-829-0515 (international or local)
707-829-0104 (fax)

We have a web page for this book, where we list errata, examples, and any additional information. You can access this page at *http://bit.ly/augmented-human*.

To comment or ask technical questions about this book, send email to *book-questions@oreilly.com*.

For more information about our books, courses, conferences, and news, see our website at *http://www.oreilly.com*.

Find us on Facebook: *http://facebook.com/oreilly*

Follow us on Twitter: *http://twitter.com/oreillymedia*

Watch us on YouTube: *http://www.youtube.com/oreillymedia*

Acknowledgements

I consider myself tremendously lucky to have an incredibly loving, supportive, kind, patient, encouraging, and inspiring family who has lifted me up throughout life, not just in the process of writing this book. My dear parents, your love and care is bountiful and you are tirelessly dedicated to our family's happiness and wellbeing. It is with great love and gratitude that this book, and everything I do, is for my family.

Thank you Caitlin Fisher for your amazing mentorship throughout my M.A., and Ph.D., and for inviting me to be a part of York University's AR Lab 12 years ago. It was life-changing! You opened my eyes to the wonder and magic of AR with the fantastical way you see the world. I am grateful and in such awe of you.

Without my editor Jeff Bleiel, whose helpful guidance, patience, and enthusiasm for the subject made for a very supportive and enjoyable writing experience, and Susan Conant and Laurel Ruma at O'Reilly, this book was but a dream. Thank you each for making it a reality and believing in my work.

Special thanks to Tom Furness, Tom Emrich, Matt Miesnieks, Soraya Darabi, Stefan Sagmeister, Jody Medich, Al Maxwell, Jonah, Dan, Mary, Sophie, Tom, Fredelle, and Martin.

And thank you, to you my reader, for picking up this book. We have the incredible opportunity and privilege to design the future; let's make it truly outstanding.

A New Wave of Reality

You are about to enter a new reality. Here, the world augments itself to you, morphing to your context, preferences, and needs. Reality becomes malleable, mutable, and highly personalized; it's all defined and driven by you. The entire world becomes instantly translatable, breaking communication barriers, and creating a new sensory awareness that makes seeing, hearing, touching, and tasting brand new. The rules of the analog world no longer apply. Wearable computers, sensors, and intelligent systems are extending our human abilities and giving us superpowers.

This is the new Augmented Reality. Are you ready?

In this book, I'll introduce you to Augmented Reality (AR), how it is evolving, where the opportunities are, and where it will go. I'll guide you to a new dimension and immersive experience medium. However, you won't need to leave your physical reality behind. The digital enters your world.

Let me explain.

This book is not about Virtual Reality (VR), but it's worth understanding how AR and VR differ.

In VR, you strap on a special headset that blocks out your view of the physical world, trading the real world for a completely computer-generated environment.

SnowWorld, developed at the University of Washington Human Interface Technology (HIT) Lab in 1996 by Hunter Hoffman and David Patterson, was the first immersive VR world designed to reduce pain in adults and children. SnowWorld was specifically developed to help burn patients during wound care. Hoff-

man explains[1] how VR helps to alleviate pain by distracting patients from their present physical reality:

> *Pain requires conscious attention. The essence of VR is the illusion users have of going inside the computer-generated environment. Being drawn into another world drains a lot of attentional resources, leaving less attention available to process pain signals.*

VR is reliant upon the illusion of being immersed in another space and time, one that is typically removed from your current reality. In AR, you remain in your physical world and the virtual enters your surroundings by way of a pair of see-through digital glasses, a smartphone, a tablet, or a wearable computer. You still see and experience the real world around you with all of your senses, it just now becomes digitally enhanced and alterable.

One early application of AR to deliver a helpful experience was Word Lens.[2] Imagine travelling to a new country where you aren't fluent in the local language. Ordering food from a menu, or reading road signs can be challenging without someone to assist you. Word Lens allows you to point your smartphone at printed text in a foreign language and translate it on the fly into the language of your choice. Suddenly, you are more deeply immersed and engaged with your surroundings via a newfound contextual understanding assisted by technology.

VR will have its dedicated uses, but AR allows us to be more deeply immersed in, and connected to, the real world—the world in which we actually spend the majority of our time and attention. As with VR, we must be cognizant of draining our "attentional resources" in AR and design experiences that do not further separate us from our surroundings or one another. We must think critically about how we will place human experience at the center of this new medium. It's not about being lost in our devices; it's about technology receding into the background so that we can engage in human moments.

What Is Augmented Reality?

The most commonly used definition of AR is a digital overlay on top of the real world, consisting of computer graphics, text, video, and audio, which is interac-

1 Hunter Hoffman, "Virtual Reality Pain Reduction," University of Washington Seattle and U.W. Harborview Burn Center.

2 The Word Lens technology was acquired by Google in 2014 and integrated into the Google Translate app. It now supports 37 languages.

tive in real time. This is experienced through a smartphone, tablet, computer, or AR eyewear equipped with software and a camera. You can use AR to point at and identify stars and planets in the night sky, or delve deeper into a museum exhibit with an interactive AR guide. AR presents the opportunity to better understand and experience our world in unprecedented ways.

We've been using the same definition of AR since 1997 when AR pioneer Ronald Azuma succinctly explained,[3] "AR allows the user to see the real world with virtual objects superimposed or composited with the real world. Therefore, AR supplements reality, rather than completely replacing it."

AR technology traditionally works by tracking a target in the real world using a camera and software on an enabled device like a smartphone. These targets can include things like an icon, an image, an object, a sound, a location, or even a person. The target input data is processed by the software and compared against a database of potentially corresponding information. If there's a match, an AR experience is triggered and content is superimposed on top of reality.

Azuma's definition states[4] that AR systems have the following three characteristics:

- Combines real and virtual
- Interactive in real time
- Registered in three dimensions (3-D)

Registration, the third characteristic, is about seamlessly aligning virtual objects into 3-D space in the real world. Without accurate registration, the illusion of virtual objects existing in the physical world is compromised; the believability is broken. So, if a virtual AR lamp appears to float above your physical desk rather than being registered directly to the table, other than believing your office is haunted, this technical glitch breaks the illusion of that lamp existing in your space. But when a shadow is added to the virtual object, it becomes even more believable because it mirrors the characteristics of your physical environment.

---

3 Ronald T. Azuma, "A Survey of Augmented Reality," *Presence: Teleoperators and Virtual Environments 6* 4 (1997): 355-385.

4 Ibid.

How AR Is Evolving

What's missing for me in this definition today, and what distinguishes the next wave of AR, is one keyword: context. Contextual information transforms the AR experience and content because it now moves from an experience that is the same for every user to one that is specific to you, your location, your interests, and your needs. Context builds on the characteristic of registration because it is registering, or compositing, relevant and meaningful data on top of the real world to create a personalized experience for you.

The success of this contextual registration in the new AR will not be about a virtual lamp looking like it sits perfectly on your physical desk (as in the 1997 definition). It will be about that lamp appearing at the appropriate moment when you perhaps need more light, or even turning itself off to indicate that it is time for you to leave work. Technical registration will be solved, and although it will continue to be important, the focus will be on delivering a meaningful and compelling experience that enhances your reality.

The process of target matching now becomes more complex because it is no longer a "hit play" process connected to a static library of things—like a photo of a dinosaur in a textbook which then triggers a 3-D model of a dinosaur displayed in AR. Today, that 3-D model and experience can be adaptive dependent on factors like how far each student has progressed in a lesson plan, and even the student's learning style. So, the next time the student returns to the AR book, the dinosaur species has changed and it integrates other topics in which she is interested. AR technology becomes a living, breathing database: an interaction in which both the triggers and the content are dynamic and can change at any moment because they adapt to your shifting contextual data to deliver timely and relevant information and experiences dictated by you and your environment.

We are well overdue to revisit what AR is and what it can become, especially with AR no longer limited to academic research facilities. AR once required highly specialized equipment, none of which was very portable. But with the number of sensors in your smartphone today, you have the power of AR in your pocket. The technology will continue to become more ubiquitous with wearable computing discreetly embedded in your clothing and glasses, and even under your skin.

Large companies like Apple, Facebook, Microsoft, Google, and Intel are paying close attention to and investing in AR's future to bring it to a mass audience. Facebook CEO Mark Zuckerberg refers to AR as "a new communication plat-

form." He writes, "One day, we believe this kind of immersive, augmented reality will become a part of daily life for billions of people."[5]

Apple CEO Tim Cook regards[6] AR "as a big idea like the smartphone." Cook says, "I think AR is that big, it's huge. I get excited because of the things that could be done that could improve a lot of lives. And be entertaining."[7] In 2017, at the annual World Wide Developers Conference (WWDC), Apple introduced ARkit, a cutting-edge platform for developing AR apps for iPhone and iPad. During the WWDC keynote, Craig Federighi, Apple's senior vice president of software engineering, referred to ARkit as "the largest AR platform in the world."[8]

AR is about augmenting the human experience and it will not advance in isolation. The real impact AR will have is when it becomes a super medium that combines other parallel emerging technologies like wearable computing, sensors, the Internet of Things (IoT), machine learning, and artificial intelligence.

The first wave of AR, which I refer to as "Overlay," was all about a digital layering on top of reality. Overlay included examples like a 3-D model of a baseball player virtually appearing on a baseball trading card, or an augmented quiz game appearing on a beer coaster. There was little to no variation if you returned to the AR experience later; it was typically the exact same content, providing not much incentive for repeat experiences. Often in this first wave, you were also required to download and print a specific image or target to trigger the AR experience.

We are entering the second wave of AR, which I call "Entryway," creating a more immersive, integrated, and interactive experience. The key difference between Overlay and Entryway (and the secret to creating meaningful AR experiences) is *you*. You are the driving force in Entryway. You are the context that defines the experience.

Unlike in Overlay, this next wave moves beyond printed targets toward a new spatial understanding and deeper intelligence of your environment. The entire world becomes a trackable target. In Entryway, we break through the limitations of overlays in the first wave, stepping into a new sensory awareness and heightened interaction with our world and each other.

---

5 Mark Zuckerberg's Facebook page, March 25, 2014.

6 David Phelan, "Apple CEO Tim Cook: As Brexit hands over UK, 'times are not really awful, there's some great things happening'," *The Independent*, February 10, 2017.

7 Ibid.

8 WWDC 2017 Keynote

New sensor-equipped AR smartphones, like the Lenovo Phab 2 Pro, and the Asus ZenFone AR, powered by Google's AR technology, Tango, are a great example of Entryway at work. Tango technology incorporates motion tracking and depth perception, enabling a device to navigate the physical world the way humans do.

As you hold the device and move it around a room, the depth-sensing camera sees what you see and is able to identify the physical boundaries and layout of your surroundings. It can recognize where walls are, where the floor is, and even where furniture is positioned. In the not-too-distant future, technology such as Tango will enable new types of everyday experiences like reading your child a bedtime story. Imagine the foot of the bed transforming into a virtual safari truck as you watch a monkey jump from the dresser onto the lamp, all while a lion sleeps soundly on top of the dresser. Your physical environment is integrated into the story world, putting you directly into the story.

An Entryway to Your Senses

Microsoft Kinect introduced a major shift in the way AR technology worked to recognize a target in the real world. Kinect was instrumental in putting you inside the AR experience because your moving body now became a trackable target. Prior to Kinect, AR targets were typically static and limited to things like printed images. This technology opened the door to more interactive experiences to better see and sense you and your actions, with the ability to even recognize your facial expressions and how you are feeling. (Chapter 2 looks at how computer vision has evolved in AR and how it's giving us new eyes with which to experience the world.)

Kinect inventor Alex Kipman (and inventor of Microsoft's AR headset HoloLens) describes[9] Kinect's impact as "a monumental shift, where we move the entire computer industry from this old world, where we have to understand technology, into this new world, where technology disappears and it starts more fundamentally understanding us." AR technology not only sees us and the environment that surrounds us, it begins to understand our activity, and responds to us. The way we interact with technology becomes more natural because the technology disappears and the experience becomes central. This is Entryway.

---

9 "How The X-Box Kinect Tracks Your Moves," *NPR*, November 19, 2010.

Entryway is all about a new level of immersion: we're reaching through the looking glass of Overlay to experience the virtual with all of our senses in a new dimensionality. Engaging the human senses beyond the visual will play a more prominent role in this next wave of AR. For example, augmented audio is often paired with visuals, but sound can be used in AR on its own without a display, or even integrated with other senses. In addition to sight and sound, we now are able to touch, smell, and taste the digital, and even create new senses (these ideas are further explored in Chapters 3, 4, and 5).

In Entryway, AR embraces a new mode of hybrid physicality and virtuality. AR imbues the physical world with digital properties, and the virtual gains a new sense of tactility. Haptics technology enables a person to experience the sensations of touch and to feel the digital using interfaces such as air pressure fields, deformable screens, and special controllers. For example, AR makes it possible to reach out and pet a virtual cat, actually feeling its fur and the vibrations of it purring.

Taste and smell are also possible in AR with devices like the "Electronic Taste Machine" and the "Scentee," both inventions of Adrian David Cheok, Professor of Pervasive Computing at City University London. The Scentee, a small device which you plug into the audio jack of your smartphone, allows you to send smell messages that release aromas. The Electronic Taste Machine uses metal sensors to trick your tongue into experiencing various tastes, ranging from sour to bitter, salty or sweet, depending on the electrical current passing through the electrode. This results in a virtual taste perception in your brain.

Cheok wants us to be able to interact with computers the way we interact in the physical world using all of our five senses. He explains:[10]

> *Imagine you're looking at your desktop or your iPhone or your laptop, everything is behind a glass, behind the window, you're either touching glass or looking through glass. But in the real world, we can open up the glass, open the window and we can touch, we can taste, we can smell.*

This next wave of AR allows us to "open up the glass" and augment the human sensorium.

The human brain can translate digitized and electrochemical signals to create meaning and even new sensory experiences. Humans currently don't see

10 "Share touch, smell and taste via the internet."

things like radio waves, X-rays, and gamma rays because we don't have the proper biological receptors. It's not that these things are unseeable; humans cannot see any of these, at least not yet, because we are not equipped with the proper sensors. AR can give us these new superpowers to not only see, but use our entire bodies to fully experience a broad range of information and data. We have the technology to engage with and know our world in extraordinary ways.

AR Across Industries

Let's take a look at a few of the industries this next wave of AR is already affecting.

AUGMENTED HEALTH

AR has made it possible for medical practitioners to interact with virtual 3-D scaled models of human anatomy. A physician now can manipulate a digital model and even 3-D print different stages of a procedure. Recent developments in haptics will enable a surgeon to one day work on a virtual brain physically, engaging in a full tactile experience before performing a real life operation.

AUGMENTED LEARNING

We can use AR today to track facial expressions to see when a student is struggling. Teachers will be able to use this technology in the near future to alter and customize content to learners. For instance, if you're taking a distance learning course or watching a lecture through your AR device, and you appear confused, the subject matter would be further explained to you. Alternatively, if you're not paying attention, you might be prompted with a question.

AUGMENTED RETAIL

AR currently enables you to see how products appear in your home, such as furniture, and how they look on your body, like a watch or article of clothing. Advancements are underway to allow consumers to not only see how products and clothing look, but also touch and feel them.

AUGMENTED WORK

AR already provides instructional repair guidance with the ability to share what you're looking at and receive real time annotation. New design processes, allowing real time remote collaboration are emerging and will change the way we work across distances. For example, an architect in Japan could be on location with a builder in Canada, interacting and fully engaged at the job site.

AUGMENTED ENTERTAINMENT

One day, you might no longer need a television: your AR headset will become your entertainment hub, full of personalized content. Whether your favorite performer appears in your home and sings to you, or you're in an open field and competing to make it through a virtual maze, new forms of digital content will be tailored to and coexist with your physical surroundings.

Today's AR: Focusing on the Human Experience

When I began working in AR 12 years ago, the primary focus of the field as a whole was on the technology; content came much later, if it even did, and it was typically an afterthought. At a time when most researchers and developers were working on registration and tracking problems in AR, I was fortunate to be a member of an extremely unique lab at York University in Toronto, Canada (led by Dr. Caitlin Fisher), where we were working on defining the future of AR storytelling. Our lab was very different from other research facilities at the time: we were based in the Faculty of Fine Arts and Department of Film, whereas most university research labs in AR were found in computer science departments. Other labs typically focused on one particular area of technical research in AR and specialized to invent and refine those techniques. Our lab, on the other hand, was centered on creating content and experiences.

We were software and hardware agnostic in our approach. The technology inspired the experiences we designed, but we didn't limit ourselves to any of the technical restrictions of AR. There were a lot of labs working on solving those problems; the area that was not being explored was content creation and the new types of experiences this technology would allow. We experimented with multiple emerging technologies, combining them in new ways to push beyond the limitations of how AR was traditionally used. If the technology didn't exist, we collaborated with engineers and scientists to create it.

In 2009, our lab developed SnapDragonAR, one of the first commercially available drag-and-drop software tools to enable nonprogrammers to build experiences and contribute to this new medium, making AR accessible to educators, artists, filmmakers, and a general audience. This created a gateway to content production for makers of all kinds. We expanded the world of AR beyond the technical realm of computer science, with innovators working in AR today continuing down this path.

AR is no longer just about the technology; it's about defining how we want to live in the real world with this new technology and how we will design experien-

ces that are meaningful and help advance humanity. The technology, awareness, and state of AR have evolved tremendously over the past decade. Now that we have all of this incredible technology, what are we going to do with it? This is our question to collectively answer as we define AR's trajectory. We are in need of leaders across business, design, and culture to help steer and implement new experiences in this rapidly rising industry. AR will radically change the way we live, work, and play.

Seeing the World Anew

We are at the beginning of a massive change in how we see and experience reality. Computer vision, machine learning, new types of cameras, sensors, and wearable devices are extending human perception in extraordinary ways. Augmented Reality (AR) is giving us new eyes.

AR's evolution as a new communications medium is rooted in the history of the moving image and early cinema. In 1929, pioneering filmmaker Dziga Vertov wrote about the power of the camera to depict a new reality, "I am a mechanical eye. I, a machine, show you the world as only I can see it." Vertov's famous film *Man with a Movie Camera* used innovative camera angles and techniques to defy the limitations of human vision.

Vertov experimented with novel vantage points (such as filming from moving vehicles like a motorcycle, to placing a camera on the train tracks while a train passed overhead). He also explored a new sense of time and space by superimposing images and speeding up and slowing down film. Vertov used the emerging technology of the mechanical camera to extend the capabilities of the human eye and create new ways of seeing. He wrote, "My path leads to the creation of a fresh perspective of the world. I decipher in a new way a world unknown to you."

Nearly a century later, Vertov's path has led us to AR revealing a new reality and understanding of our world. The camera plays a central role in how AR technology traditionally works: a camera is paired with computer vision to scan and decipher our physical surroundings. AR previously relied heavily on fiducial markers (black and white geometric patterns) or images to augment two-dimensional (2-D) surfaces, such as a print magazine.

The real world, however, is not flat; we experience it in three-dimensional (3-D) space. Unlike 2-D fiducial markers or images, 3-D depth-sensing cameras are being used in AR to recognize, map, and understand our spatial surroundings. These 3-D depth-sensing cameras, such as the Microsoft Kinect camera and Intel's RealSense camera, are replacing the use of fiducial markers and images to change the way computers see, translate, and augment 3-D environments.

Vertov's work explored how the camera as a mechanical eye could defy the limits of human vision. He presented novel perspectives depicting what it would be like if a human could see like a camera. Depth-sensing cameras like Kinect and RealSense present the opposite: what if a camera and computer could see like a human? AR technology is beginning to mimic the design of human sensibilities allowing us to see in completely new ways.

You Are the Controller

When introduced in 2010, Kinect changed the way we experienced AR. Kinect's tag line was "You are the controller." By simply moving your body as you naturally would, you triggered and directed the AR experience.

Prior to Kinect, for AR to appear on your body, you would need to cover yourself in 2-D fiducial markers, have an image printed on your clothing, or get an AR tattoo. But with Kinect, the experience instantly became more immersive because there was no barrier between you and the augmentation; it was you. Standing in front of a screen powered by Kinect, you could see and interact with a transformed version of yourself, as though standing in front of a magical digital mirror. The augmentations followed your movement and responded to your gestures, creating an experience unique to you.

Artists immediately embraced Kinect as a creative tool to build new types of interactive experiences. Chris Milk's "The Treachery of Sanctuary" (2012), is a beautiful example of Kinect used in an art installation. You are invited to stand in front of a series of three interactive panels that represent the creative process through birth, death, and regeneration. Your body is mirrored back to you as a dark shadow with different transformations occurring in each panel. In the first panel, your body disintegrates into rising birds. As you move into the second panel, these same birds swoop in to assail you. In the third and final panel, your body sprouts giant wings, and by flapping your arms, your form takes flight, rising off the ground and ascending into the sky.

Milk writes in an artist statement:[1]

> *What is interesting to me is the two-way conversation between the work and the viewer. The participant is an active character in the content and concept of the piece, and while the technology allows that interactivity, the emphasis is on the experience, on transcending past the enabling innovation to the spiritual immersion.*

Part of Kinect's magic is that the technology becomes invisible because it is easy to use: you stand in front of it and move around. The experience is reactive to you; your body movements trigger what happens. The technology enables the experience, but without you, there is no content. The technology recedes into the background and you become the focus, quite literally.

Observing Movement and Predicting Activities

Kinect uses a depth-sensing camera to see the world in three dimensions. It works by projecting a pattern of infrared light points onto a room and then measuring how long it takes the light from each of those points to reflect back to the camera's sensor chip. Software processes the data to identify any human shapes that might be in view, like heads or limbs. Kinect uses a skeletal model that breaks down the human body into multiple segments and joints. Programmed with more than 200 poses, the software understands how a human body moves and is able to predict what movement your body is likely to make next.

Prediction is an important aspect of human perception that we use extensively in our daily activities to interact with our surroundings. Jeff Hawkins, the founder of Palm computing—the company that gave us the first handheld computer—and author of the book *On Intelligence* (Times Books, 2004), describes the human brain as a memory system that stores and plays back experiences to help us predict what will happen next.

Hawkins points out that the human brain is making constant predictions about what is going to happen in our environment. We experience the world through a sequence of patterns that we store and recall, which we then use to match up against reality to anticipate what will happen next.

Using Kinect, researchers at Cornell University's Personal Robotics Lab programmed a robot to anticipate human actions, and assist in tasks like pouring a

1 "The Treachery of Sanctuary."

drink or opening a refrigerator door. The robot observes your body movements to detect what action is currently taking place. It accesses a video database of about 120 household activities (ranging from brushing your teeth, to eating, to putting food in the microwave) to predict what movement you will make next. The robot then plans ahead to assist you in that task.

Building a 3-D Map with SLAM Technology

For a robot to move through an environment and perform such activities, it needs to be able to create a map of its surroundings and understand its location within it. Roboticists have developed Simultaneous Localization and Mapping (SLAM) technology to accomplish this task. Prior to SLAM the sensors required to build that map have traditionally been expensive and bulky. Kinect introduced an affordable and lightweight solution. Videos of Kinect-enabled robots appeared on YouTube within weeks of Kinect's release. The robots ranged from a quadrotor flying autonomously around a room, to a robot capable of navigating rubble to find earthquake survivors.

Google's self-driving car also uses SLAM technology with its own camera and sensors. The car processes both map and sensor data to determine its location and detects objects around it based on their size, shape, and movement. Software predicts what the objects might do next and the car performs a responsive action such as yielding to a pedestrian crossing the street.

SLAM is not limited to autonomous vehicles, robots, or drones; humans can use it to map their environment, too. MIT developed one of the first examples of a wearable SLAM device for human users. The system was initially designed for emergency personnel like first responders who enter unknown territory. With a Kinect camera worn on the chest, a digital 3-D map is built in real time as the user moves through the environment. Specific locations can be annotated using a handheld pushbutton. The map can be shared and immediately transferred wirelessly to an offsite commander.

SLAM also makes possible new forms of game play. *Ball Invasion*, created in 2011 by 13th Lab in Stockholm, Sweden, is an early example of integrating SLAM into an AR game. Holding your iPad in front of you, you see your physical surroundings filled with virtual targets to shoot and chase. What made *Ball Invasion* unique is that the virtual elements interact with your physical world: virtual bullets bounce off the wall in front of you, and virtual invading balls hide behind your furniture. As you play the game and move the iPad's camera around, you are building a real time 3-D map of the environment to enable these interactions.

In 2012, 13th Lab released PointCloud, an iOS Software Development Kit (SDK) with SLAM 3-D technology for app developers. 13th Lab was acquired by VR technology company Oculus in 2014.

Today, SLAM is one of the underlying technologies behind Google's Tango AR computing platform. In 2015, tablet development kits for Tango became available to professional developers first, with Tango-enabled smartphones introduced later in 2016 (the Lenovo Phab 2 Pro) and 2017 (the Asus ZenFone AR). Tango makes possible experiences such as precise navigation without GPS, windows into virtual 3-D worlds, measuring spaces in real time, and games that know where they are in a room and what's around them. Google describes[2] the goal of Tango as giving "mobile devices a human-scale understanding of space and motion."

Our smartphones are already an extension of ourselves, and with advancements like Tango, smartphones are beginning to see, learn, and understand the world like we do. This will give way to new types of interactions in which the virtual is seamlessly mapped to our physical reality and is contextually relevant, creating a deeper sense of immersion. The lines between the virtual and real will blur even more. Technology will not only understand our surroundings, but perhaps help navigate us through our daily lives with a new-found intelligence and awareness.

Helping the Blind to See

If we can bring vision to computers and tablets, why not use that same technology to help people see? Rajiv Mongia, director of the Intel RealSense Interaction Design Group, and his team have developed a portable prototype of a wearable device that uses RealSense 3-D camera technology to help people who are vision-impaired gain a better sense of their surroundings.

Demonstrated live on stage at the 2015 International Consumer Electronics Show (CES) in Las Vegas, the RealSense Spatial Awareness Wearable consists of a vest fitted with a computer that connects wirelessly to eight thumb-sized vibrating sensors worn across the chest, torso, and near the ankles of each leg. It works by seeing depth information to sense the environment around the wearer. Feedback is sent to the wearer through haptic technology that uses vibration motors for tactile feedback.

2 Google Tango

The vibration sensors are comparable to the vibration mode on your phone with the intensity of the vibration being proportional to how close an object is to you. If an object is very close, the vibration is stronger, and if it's farther away, it is lower.

Darryl Adams, a technical project manager at Intel, has been testing the system. Adams was diagnosed with retinitis pigmentosa 30 years ago and says the technology allows him to make the most of the vision he does have by augmenting his peripheral vision with the sensation of touch.

> *For me, there is tremendous value in the ability to recognize when change occurs in my periphery. If I am standing still and I feel a vibration, I am instantly able to turn in the general direction to see what has changed. This would typically be somebody approaching me, so in this case I can greet them, or at least acknowledge they are there. Without the technology, I typically miss this type of change in my social space so it can often be a bit awkward.*

The system was tested on three wearers, each with very different needs and levels of vision, from low vision to fully blind. Mongia and his team are working on making the system scalable with modular hardware components to allow users to select the combination of sensors and haptic output that best suits their specific situation.

Adams would like to see the software become context-aware so that the system can respond to the wearer's needs in any given scenario. He thinks this technology could evolve to include features like facial recognition or eye tracking. This way, the wearer can be alerted when someone is looking at them rather than just knowing there is a person nearby.

Artificial Intelligence (AI) could further be integrated to provide wearable systems with a better understanding of the wearer's context. Methods like machine learning can help give computers some of the abilities of a human brain, enabling computer programs to learn to perform new tasks when exposed to new data, without being explicitly programmed for those tasks.

Teaching a Computer to See with Machine Learning

OrCam, a device designed for the visually impaired, uses machine learning to help wearers interpret and better interact with their physical surroundings. The device can read text and recognize things like faces, products, and paper currency.

The OrCam device features a camera that clips on to a pair of glasses and continuously scans the wearer's field of view. The camera is connected by a thin cable to a portable computer that fits in your pocket. Instead of vibration sensors (like in the RealSense Spatial Awareness Wearable), OrCam uses audio. A bone-conduction speaker transmits sound to the wearer as it reads aloud objects, words, or people's names.

With OrCam, the wearer shows the device what he is interested in by pointing. "Point at a book, the device will read it," says Yonatan Wexler,[3] head of research and development at OrCam. "Move your finger along a phone bill, and the device will read the lines letting you figure out who it is from and the amount due." To teach the system to read, it is repeatedly shown millions of examples, so the algorithms focus on relevant and reliable patterns.

Wexler says there is no need to point when identifying people and faces. "The device will tell you when your friend is approaching you. It takes about ten seconds to teach the device to recognize a person," he says. "All it takes is having that person look at you and then stating their name." OrCam takes a photo of the person and stores it within the system's memory. The next time the camera views the person, the device will recognize that person, and even identify them by name.

OrCam uses machine learning to recognize faces. The research and development team had to provide OrCam with hundreds of thousands of images of all different kinds of faces in order to teach OrCam's program how to recognize an individual face. When a user is wearing OrCam, the program sorts through all images, rejecting ones that are not a match, until only the one matching picture remains. In a matter of moments, this process of face recognition occurs each time someone wearing OrCam encounters someone they have taken a picture of with their device.

Training the Brain to See with Sound

OrCam is trained to see your world and provide an oral translation of your immediate surroundings. A different approach is taken by vision technologies like the vOICe and EyeMusic. Instead of using machine learning to tell the wearer what she is looking at, these technologies explore how the human brain can be trained to see with other senses, effectively learning how to see with sound.

---

3 Helen Papagiannis, "Augmented Reality Applications: Helping the Blind to See," *iQ*, February 10, 2015.

Neuroscientist Amir Amedi asks, "What if we found a way to deliver the visual information in the brain of individuals who are visually impaired, bypassing the problems in their eyes?" Brain imaging studies conducted by Amedi and his team show that when people who have been blind since birth use systems like the vOICe and EyeMusic to "see," they activate the same category-dependent processing areas of the brain as people who are sighted. However, instead of traveling through the visual cortex, the signal enters the brain through the auditory cortex and is diverted to the proper spot of the brain.

The vOICe system (OIC = "Oh, I See") translates images from a camera into audio signals to assist people who are congenitally blind to see. Developed by Peter Meijer, the vOICe consists of a pair of sunglasses with a small camera that is connected to a computer and a pair of headphones. (The system also can be used on a smartphone by downloading the software and using the phone's built-in camera.)

The vOICe software makes your surroundings into a "soundscape." The camera continuously scans the environment from left to right, converting each pixel into a beep: the frequency represents its vertical position, and the volume of each beep represents the brightness of the pixel. Brighter objects are louder, and frequency indicates whether an object is high or low.

Amedi and his colleagues have trained people who were born blind to "see" using the vOICe and EyeMusic, a more recent app developed by Amedi that additionally assigns different pitches to colors. Different types of instruments are used to convey colors. For example, blue is signified by a trumpet, red by chords played on an organ, and yellow by a violin. White is represented by human voices, and black is silence.

Amedi says training one's brain to learn to see this way takes about 70 hours. Users are taught how to identify broad categories of objects, including faces, bodies, and landscapes. Each is processed in the visual cortex of the brain. "Everyone thinks that the brain organizes according to the senses, but our research suggests this is not the case," says Amedi.[4] "The human brain is more flexible than we thought."

Research and inventions like Amedi's and Meijer's show us that the traditional definition of what it means to see is changing. It will continue to change as both computers and the human brain are learning to see in new ways together.

---

4 Roni Jacobson, "App Helps the Blind 'See' With Their Ears," *National Geographic*, April 5, 2014.

Choose Your Own Reality

The ability to see and interpret our surroundings with the assistance of computer vision also makes it possible to filter our reality and selectively see, or unsee, the world around us. This includes the possibility of removing things from our reality that we do not wish to see.

Black Mirror, a popular television series satirizing modern technology, imagined the ability to block people in real life with the press of a button in the episode "White Christmas" (2014). Instead of seeing the person you have blocked, you now see a white space in the shape of a person, and hear muffled sounds, with the blocked person seeing the same for you. In 2010, Japanese developer Takayuki Fukatsu built a demonstration not too different from the technology presented in the episode of *Black Mirror*. Using Kinect and OpenFrameworks, Fukatsu's Optical Camouflage shows a human figure blending in with his background to become invisible.

Dr. Steve Mann is a professor of electrical engineering and computer science at the University of Toronto and is referred to as "the father of wearable computing." Mann defined the term "Mediated Reality" in the 1990s. He says,[5] "Mediated Reality differs from virtual reality (or augmented reality) in the sense that it allows us to filter out things we do not wish to have thrust upon us against our will." For Mann, wearable computing devices provide the user with "a self-created personal space." Mann has used Mediated Reality to substitute personal notes and directions in place of advertisements.

New media artist Julian Oliver credits Mann's work as inspiration for "The Artvertiser," a Mediated Reality project initiated in 2008, which he developed in collaboration with Damian Stewart and Arturo Castro. The Artvertiser is a software platform that replaces billboard advertisements with art in real time. It works by teaching computers to recognize advertisements, which are then transformed into a virtual canvas upon which artists can exhibit images or video. The artwork is viewed through a handheld device that looks like a pair of binoculars.

Rather than referring to this as a form of AR technology, Oliver considers The Artvertiser to be an example of "Improved Reality." He describes the project as a reclaiming of our public spaces from "read-only" to "read-write" platforms. The Artvertiser applies a subversive approach to reveal and temporarily intercept environments that are dominated by advertising.

---

5 Steve Mann, "Mediated Reality: University of Toronto RWM Project," *Linux Journal*, March 1, 1999.

The "Brand Killer" (2015) is a contemporary project that builds upon Mann's and Oliver's work. Brand Killer was created by a group of University of Pennsylvania students, Tom Catullo, Alex Crits-Christoph, Jonathan Dubin, and Reed Rosenbluth, to blur ads in real time for its wearer. The students question, "What if we lived in a world where consumers were blind to the excesses of corporate branding?" Brand Killer is a custom-built head-mounted display that uses openCV image processing to recognize and block brands and logos from the user's point of view in real time. It's "AdBlock for Real Life," they state.

We already mediate our reality while we're on the internet by blocking ads and even people with whom we no longer want to interact. Beyond advertising and other people, what else will we choose to remove or block from our vision with Mediated Reality?

As we design the future of AR, we will need to consider if digitally filtering, mediating, and substituting content to one's choosing will enhance our reality or separate us from the world and one another. It is my hope that these new technologies will be used in ways to support human interaction, connection, and communication, and even build empathy.

Although we are often inclined to erase things from reality that we do not want to see, such as homelessness, poverty, and sickness, there are things that we, as society, must actively address. Mediated Reality has the potential to foster a culture of avoidance and even ignorance. We should not turn a blind eye to the realities of reality.

The positive side of Mediated Reality is that it could be used as a way to provide focus. This technology has the possibility to create a future with less distractions that leads to more human-to-human moments. We are already bombarded by technology and notifications; what if Mediated Reality provided an easy way to completely switch off distractions temporarily?

Another critical question is who will be authoring this new reality? Will it be individuals, corporations, or groups of people? Whose Mediated Reality will we be privy to and what visual filters or tools for interception will come to exist? To use Oliver's terms, will we be part of a read-write environment, or read-only?

In the same way that the internet is read-write, I believe AR, with Mediated Reality a part of it, will be, too. Tim Berners-Lee, inventor of the World Wide Web, envisioned the internet as a place to share experiences in new and powerful ways. "The original thing I wanted to do was make it a collaborative medium, a

place where we can all meet and read and write," he says.[6] The internet reframed the way we share and consume information and AR has the power to do this, as well.

With examples such as enabling the visually impaired to gain a form of sight, artists imagining new interactive experiences, and robots assisting humans in daily life, AR presents a new way of perceiving the world. AR has the ability to improve people's lives and inspire life changing ways of engaging with our surroundings and one another.

If we replace the word "machine" with the word "human" in Vertov's sentiment at the beginning of this chapter, "I, a machine, show you the world as only I can see it," we get the richness of what the internet enables: a global collection of shared stories of human experiences and perspectives. To have a positive impact on society and contribute to humanity in a meaningful way, AR will need to find ways to mirror the original vision for the World Wide Web to largely be inclusive, not exclusive.

6 Andy Carvin, "Tim Berners-Lee: Weaving a Semantic Web," February 1, 2005.