A New Way To See

Transcript from TEDx, 2016

Stephen Hicks, PhD

How much do you see of the world around you?

Is it colourful and rich? Are objects around you sharp and clear?

We feel like we take all this information in through our eyes. Like they are a high definition video camera. However, in the few seconds that you've been listening to me, you've really only seen a tiny fraction of the visual world.

The area of the eye that takes in sharp detail, and almost all of the colours, is called the fovea and it covers as much visual space as your thumb at arm's reach. Within that tiny window lie almost all your colour detecting cells, and your ability to resolve fine lines.

But what I see when I look around me is a clear, and sharp understanding of all the objects in it. However, in reality, everything outside of the fovea is blurry.

Take reading for example.

mountain .

To read a passage, we move our eyes rapidly across the words and phrases on a page. However if you look just to the side of a word (say at this dot on the right) and keep your eyes fixed on the dot and mentally look at the word to the left, it's practically unreadable.

So how come the world doesn't seem blurry everywhere except right in the middle?
Neuroscientists and philosophers argue that we construct or extrapolate most of the visual world, based on our memories, our predictions, and a very vague amount of vision.

How does this work?

When we come into a new room for the first time - we instinctively note the locations of the walls and the major objects within the room. Through this we form a spatial memory of the room which we use for seeing and navigation. Once we have learned a space, we use our eyes to spot changes or study dynamic and interesting things, like faces and text. Our spatial memory provides the general framework on which new

visual information can reside, and hence it's this memory component that accounts for most, nearly all, of what we see at any one time.

For many of the world's 40 million or so blind people, this is a serious problem. What happens when you aren't able to see enough of the visual world to form these types of spatial maps? What does the world around you feel like? Many people compensate by listening for the echoes of a room to get a feel for its dimensions, or look into any remaining areas of sight for clues about objects around them. But the concentration required to perform these acts can be exhausting and the results are often an impoverished sense of space around you.

What I sought to do was to find a way of giving people a better sense of their space around them, so that they can form richer mental maps and essentially see much better.

The majority of blind people have some remaining vision. This is called residual vision and it can take many forms such as: just peripheral vision in the case of macular degeneration, or just a few degrees in the centre, in the case of tunnel vision (and many other forms besides).

What I wanted to do was find ways of putting very simple, relevant spatial information into those remaining areas, to help the brain fuse this visual information into a more complete and richer sense of space.

The problems we faced were:
What types of things does a person need to see from day to day.
And how can we make that visible to someone who is nearly blind?

To solve this, we have used two very exciting modern technologies: augmented reality and computer vision. And this project has really been about the merging of computer vision, to augment and advance human vision.

We all have pretty sophisticated computer vision systems in our pockets and handbags. The software linked to most modern phones is able to detect barcodes and faces, sometimes even text. But it wont detect chairs, cars and trees and all the myriad of objects that a person might interact with during the course of a day. What we needed for our project, was a more general way of looking at the world.

Very often, the most important things are the ones that are right in front of you. We reasoned that if we could find those, we could make a high contrast visual display that might be able to help blind people see.

The brain is very good at guessing what an object is. With just a few well placed lines, we can see a whole expressive person

¯_(ツ)_/¯

This is lucky, because our first prototype was very simple indeed.

We began by showing nearby objects as bright regions of light. We made a test system using a depth camera and an array of LEDs and attached a pair of ski goggles. As you see in this video, that's enough information for you to infer that there is a person in front of you, and that they're waving. The trick is to remove all the information from the background so that a person can focus on just one or two items that are right in front.

We were able to test this with a range of blind and partially sighted people and nervously waited to see if they would be able to make anything out of this. The first person we had such restricted vision that she could normally only see people in very bright light at arm's reach. After a few minutes of trying our goggles she could see people waving at 4 metres away. Later we tested a very sight impaired gentleman, and he was able to look down and see his guide dog and said "There you are girl."

For our small team of researchers, this was incredible. It gave us the sense that it was possible to augment vision for people who were all but blind. What we needed to do next was work on the portability and look of the system, and to dig deeper into computer vision to find new ways of enhancing and detecting objects in the world.

Skip forward a few years and you have our Mark 5

(holds up prototype)

This is a see-through display that allows a person to view the world as they normally would, and inside the lenses, is a small transparent digital display on which we can project an enhanced image of the world.

We have continued to improve our algorithms so that now we can present bright high contrast lines and images that represent objects in more detail. As before we remove the background so that a person can focus on the objects in front.

This increased level of detail has allowed some people to see many more objects simultaneously, and also incredibly, to see faces and expressions better than they have for many years.

We built this particular system so that we could see what people did with them when they took them home. The results have been remarkable.

We have had people say to us that the glasses made them feel safe, that they could control their own space.

Many people felt more confident and secure with the glasses with one of our long-term participants commenting that, "**The last few weeks have changed my social life completely. I have been out more than I have in the last few years.**"

And remarkably, many people found that they could use the glasses to see faces and expressions better, with one father saying that he "**hadn't seen [his] daughters' faces for years, and being able to sit around the dinner table and see them was phenomenal."**

It's a long way from perfect though, and most people with advanced macular degeneration haven't enough sight to see display which we hope to address in future versions.

And of course, many people commented on the size and bulk of the glasses. One of my favourite quotes was,

"They look awful!"

So we are really on our way to improving people's sight in a rich and meaningful way.

The journey hasn't finished though. Obviously we are working on sleek and stylish glasses. We are experimenting with new types of displays that may be able to help a wider range of people. We are also stretching our scope of computer vision, allowing the glasses to provide more than increased contrast, but

also an intelligent extension to your sense of vision. Using our groups experience in artificial intelligence, we are testing systems that can automatically detect and highlight complex objects in any environment.

Take this photo of my daughter holding her bike. The lighting is terrible and it's hard to see anything. When we run it through a neural net classifier, we can automatically find and highlight the person and the bike.

This type of extended cognition has the potential to boost functional vision in a wide range of situations for blind and partially sighted people. The technology carries with it, the potential to learn about the objects and spaces that are specific to an individual, thereby becoming an important extension and interface through which people see and experience a richer, safer and more meaningful world.

These types of intelligent detection and enhancements could eventually be used for everyone, even those without vision impairments. You can imagine a scenarios where it might be useful to highlight people that you know in a crowd, or to easily find items in a supermarket based on some new recipe that you are trying out.

And a range of new experiences that we are only beginning to imagine. Such as the detection and representation of non-visible objects like x-rays and magnetic fields, or addition of information around a museum or gallery, or even the sharing of 3D space with friends and loved ones far away, or lost in time.

The world of augmented vision is only just beginning. It makes me extremely happy that people with low vision are going to be the ones who benefit from this first. Too often new technologies are designed for the masses, and people with difficulties have to wait years before someone will make adaptations for them. In this case, visually impaired people will be the first, and it’s their experiences that will end up shaping everyone else’s.