Professor Ted Adelson is a John and Dorothy Wilson Professor of Vision Science at MIT in the Department of Brain and Cognitive Sciences and the Computer Science and Artificial Intelligence Laboratory. He is a member of the National Academy of Sciences and a fellow of the American Academy of Arts and Sciences.

Ted has published widely in areas of human vision, computer vision, and computer graphics. His current research focuses on artificial touch sensing for robotics. Ted is well known for his contributions to multiscale image representation in basic concepts in early vision, such as motion energy and steerable filters. His work on neural mechanisms or motion perception has won awards by the IEEE and other organizations. He has done pioneering work on the problems of material perception in human and machine vision.

Professor Adelson has recently developed the novel technology for artificial touch sensing called GelSight, which converts touch to images and which enables robots to have tactile sensitivity exceeding that of human skin. And we'll talk about GelSight in this podcast. Ted, thanks for your time today.

Ted: Nice to be here.

Steve: Can we start by having you explain, broadly, your research vision and some of the aspirations of your work?

Ted: All right, these days, I'm working on robotic touch. That is I'm trying to give robots a sense of touch by giving them fingers that are sensitive and soft, like human fingers, and the goal is to give them all the capabilities that we'd like robots to have and which, for the most part, they don't have yet.

Manipulation is an area that's progressing only slowly with robots. And one of the key things is they need to have good fingers. So we're trying to make fingers that can match the capabilities of human fingers.

Steve: That's exciting. So can you tell us a little bit about what inspired you to conduct this research?

Ted: Well, I am, originally, trained as a vision person. Actually, my original training is in human perception. And, from there, I got interested in computer vision. And, for many years, vision was my thing.

And then, when I had children, I discovered that children have really fascinating abilities to use their fingers and to pick things up and manipulate them, even from when they're very small. And I just got fascinated by human touch and, in particular, fascinated by the kind of signals that I saw my children picking up, even when they were little.

So I decided to use what I know about vision as a way of building a touch sensor. So the fingers we build little cameras inside, and they look at the skin from the inside. And they can see how their skin is being deformed, and that gives us the touch signal that we can process that using image-processing techniques, but it gives us an extremely high-resolution signal that is full of really great information.

Steve: That's fascinating. So how are computer vision and touch sensors the same?

Well, they're the same in the sense that they're both looking at a signal that is distributed across space and evolving in time, right? So you've got-- with vision, you've got pixels that are changing over time, and you've got a spatial array of them. And, touch on a fingertip, you've got a similar thing, a spatial array of touch pixels that are changing in time.

But touch is much more interactive than vision because, in order to get-- in order to do anything in the world, you are touching things. You are reaching out. You're manipulating them. And so the signal on your hand is something that's, partly, coming from the world and, partly, coming from the control you're exerting on the world. So touch is, in many ways, much more complicated than that vision is.

Steve: Is that more advantageous in terms of speed, compute cost, or accuracy?

Ted: So, the touch signal, we would say it's a lower bandwidth. That is the raw amount of data coming from the finger is less than what's coming through the visual system. But you've got a lot of fingers actually, and you've got to coordinate that with all the control signals that are being used to control the movements of your fingers and your limbs.

So, in the end, whether you're doing a vision system or a touch-based system, you're sort of pushing the limits on what the technology can provide at this point. You really end up running both of them at a very high bandwidth.

Steve: And can you teach a robot to look at something and predict what it will feel like and vise versa?

Well, that's an interesting question because that's a striking capability that humans have, the ability to use your vision to look at something and predict what's going to happen when you touch it. You can predict-- let's say if you're looking at a shirt or a piece of food or something that you're going to reach out and touch it, you're predicting what's going to happen when you touch it, how it's going to deform, how it's going to feel, how it's going to behave, as you manipulate it.

And that's an amazing ability that humans have. And, until recently, it was not something that you could really imagine giving to robots, but, because of the advances in machine learning, in particular, deep learning, it's now possible to teach robots very high-level, complex relationships between the visual signal and the touch signal, both the expected touch signal and the touch signal-- and I should say the motor signals, as they're evolving. So it's becoming possible really to give robots the ability to connect the visual and the tactile information.

Steve: And did you use supervised learning in addition to deep learning to create better sensing?

Ted: So supervised learning can be-- well, both supervised learning and reinforcement learning are used heavily in robotics. And both of those can be used with deep learning techniques these days. So really learning has really revolutionized everything we do and made everything much more powerful.

Steve: And can you tell us a little bit about GelSight and the promise that it holds for robotics? So GelSight is one touch technology. It's not the only touch technology. It's the one that we're pursuing.

Ted: The way it works is there's a camera inside the finger. The fingertip is made of a soft elastomer covered with a membrane. And, by looking at what's happening to the membrane from the inside, we can see exactly what's happening in terms of the mechanical interaction between the finger and the world.

It's an attractive technology because, rather than being based on some advanced, new sort of magic nano something or other, it's just based on very straightforward stuff like cameras and elastomers, which are well known and well understood. And these days, cameras are so cheap they're almost free. And, because of cell phones, cameras are really small and really capable.

So we can take these simple, well-established technologies and repurpose them for use in this other technology, namely, touch sensing. And that gives us a way of getting a very capable sensor, which is also soft, like a human finger, and which has sensitivity that's actually better than a human finger. So we think this is a very exciting way to go in terms of giving new capabilities to robots.

Steve: So, Ted, is GelSight a product? Is it a technology, an API? Is it open source? Can you just explain a little bit more of what it is?

All right, GelSight is a particular approach to touch sensing that involves-- as I've said, it involves the use of a camera inside of a finger. There's a whole family of devices we can make using this technology. And we're publishing descriptions of how they work and making it available to the research community.

It's also a company. There's a company called GelSight, Inc. that's spun out of my lab that's, right now, not making robot devices. It's making metrology devices, but, in the future, we'll make robot devices. And there's some MIT-owned IP that's associated with the technology that's available in the usual ways that IP is available to those who are interested.

Steve: So a robot manufacturer or someone who is doing arms, actuators could take this technology and incorporate it and productize it?

Yeah, hopefully, what will happen is that, between the work in my lab and work by companies, it'll become products, starting with fingers, but it's actually-- it's difficult to make a finger that's very useful by itself. You really need to have a finger that's part of an entire system, part of a hand. And a hand has to be part of a robot, and the whole thing has to be bundled up into a system. So, in the long run, it'll just sort of disappear into a manipulation system, but it'll be a piece of it.

Steve: I see. And what industries do you think this technology will most impact?

Well, if you ask what could touch sensing be good for or, more particularly, what could good fingers be good for, basically, anything that people can do is an area where robots could do it as well. So that could be in manufacturing. That could be in food preparation. Logistics is a big area where robots are being used these days and health care, eldercare, domestic robots, robots in your house.

Basically, robots can be everywhere, doing everything. And, if they have good fingers and good hands, then they're much more capable. And one area that's of particular interest is in the areas where robots are working close by with humans. They could be working on the factory. They could be working in the home, or they could be assisting somebody, an older person getting dressed.

For these applications you need to have the capability of being-- of having good sensitivity, good touch sensitivity, to control the manipulation. You also need safety. And, again, touch is another area where-- well, which is going to be really important for safety in order to give the robot feedback to knowing what they're touching, how hard they're touching it, and is it safe the way they're touching it.

Steve: I see. Can you explain a little bit about sort of the benefits of you working with the CSAIL Alliance program and industry partners? Have you teamed up with any industry partners to advance your research?

Yes, well, so, throughout my career, I've worked both with basic research funding, like from NSF, and with funding from various industry sources. Right now, actually, the majority of my funding is coming from the Toyota Research Institute, and I'm talking to some people in CSAIL Alliances about funding possibilities.

I think that working at the boundary between basic research and applied research is really inspiring. If you just do basic research, it's sort of easy to get lost in your own little world and lose track of what the real problems are. And so, as soon as you start working on real-world problems that are brought to you by people in industry, it sort of focuses your mind, and it opens your eyes up to things you hadn't thought of.

So I think it's a terrific way of balancing the real-world information from the theoretical information. And CSAIL is a great place to do that, and the CSAIL Alliances program is a great way of fostering that kind of interaction.

Steve: Great, well, we think so too. Is there anything else you'd like to tell our listeners about sort of where you see this field going in the future?

Well, robotics is really a terrifically exciting place to be working right now because you can just begin to see the potential, but you can also see there's so much more to be done. And, basically, I think you could say in computer science, generally, a lot of computer science has been spectacularly successful working in the world of bits, the world of information.

But if you want to be working in the physical world where you're causing changes to happen, you're manipulating things, you're driving cars, you're working-- you're helping people in the physical world, you really need to work on developing great robots. And we're just beginning that, but you can see that it's really taking off.

Steve: Excellent. Can you elaborate a little bit more on how you're using deep learning for touch sensing, high-resolution sensing?

Deep learning is-- it's revolutionary. It's the most revolutionary thing I've seen in my career. And it's not a field that I ever worked in. So I'm not-- I can't claim any credit for it. But it has just deeply changed everything we do.

I think deep learning is-- or machine learning, more generally, today, for the grad students that are being trained today, it's almost like calculus. It's just a tool that everybody uses, even if you don't study it, even if you're not a theoretician in it. It has developed into a family of tools that are just so powerful that it allows you to do all kinds of stuff you couldn't before.

In the case of touch and robotics, it was just a few years ago-- let's say more than 5 or 10 years ago. If you got a touch signal coming out of your sensor, it's very hard to convert the raw data into something useful. It's very hard to figure out how you're going to use all those signals coming out of this physical device into a decision about how to move or how strong something is or whether it's slipping or something like that because those relationships between what the raw data says and what you want to know, in terms of a decision, those are very complicated relationships.

But, with deep learning, it's become possible to do those very quickly. So, something that might have taken months of hard work to figure out in the past, you can do in a few days. And you don't have to be a world's expert at machine learning because people have encapsulated this knowledge into open-source routines that you can just use. So it's really revolutionary.

Steve: Can you talk a little bit about artificial intelligence and perception? Is there-- what is the state of the art as far as perception and use of artificial intelligence with perception?

So, for me, the most important thing about intelligence, whether natural intelligence or artificial intelligence, is that the human or the robot is constantly making a model of the world. So you're making decisions about how to move your body, what to do, what's out there. You're making all these decisions based on an internal model that you're building based on the data you're gathering from the world.

And so it's this ability to capture data from the world and infer the nature of what's out there in the world and what's happening in the world and what's going to happen when you take an action. That's what's critical to intelligence. And so that's an area where people are really making a lot of progress.

In the case of manipulation, we're gathering data from our fingers that tell us about what we're touching and how things are moving and changing as we touch them. And we're using that information, along with other data, like visual data, to infer about the state of the world. And that allows us to change the state of the world by executing motor actions.

Steve: Well, that's all the time we have today. Ted, thank you very much for joining us. Have a great day.

My pleasure, thank you. Bye-bye.

[MUSIC PLAYING]