>> Mark Eichorn: Good morning, everyone. Could I get your attention, please? Thank you all for coming today. We'll be welcoming Chairman Leibowitz in a moment, but I just wanted to give some administrative logistics announcements for starters. First, if you go outside the building without an FTC badge, you'll have to go back through the magnetometer and the X-ray machine before you come back in. Secondly, if there's a fire, please go outside the building. [Light laughter] When you go to the Georgetown Law Library and Georgetown University, right outside the main entrance -- so, if you take a right and go across the street, there's a staging area where you can check in, and that way the fire folks will know that everyone's out of the building. If it's safer to remain inside the building in case of certain emergencies, then you'll be told where to go inside the building. Thank you. If you spot suspicious activity, please alert security. And finally, this event is open to the public and is being Webcast and may be recorded. Finally, if you're on Twitter and tweeting about the workshop, please use the hashtag #ftcpriv. Now it is my honor and pleasure to introduce chairman Jon Leibowitz. Chairman Leibowitz joined the Commission in 2004 and was named as Chairman in 2009 by President Obama. Throughout his service on the Commission, Chairman Leibowitz has made consumer privacy a top priority. He's strongly supported the agency's efforts to make Do Not Call registrations permanent, to get spyware off consumers' computers, to encourage self-regulatory efforts, to give consumers notice and an opportunity to control collection and use of data about them, and to bring actions against companies like Twitter, Google, and Facebook. He's also led our fundamental reexamination of how the agency approaches privacy issues. He's pushed the agency not just to address problems that occurred the past, but to think about how changes in the marketplace or in technology will affect consumers in the future. So it's particularly apt that he would kick us off today as we examine the development of facial recognition technology. With that, Chairman Leibowitz. [Applause]

>> Jon Leibowitz: Thank you, Mark, for the kind introduction. I've always had an interest in privacy, but one of the things that I was fortunate to have happen to me when I first came to the Commission was Mark Eichorn became my attorney adviser. And he just inculcated sort of the values and the importance of privacy. You were sort of a mentor to me, actually, Mark, I want to say. [Laughter] And of course he coined the phrase "Technology is a peculiar thing -- it brings

great gifts with one hand, and it stabs you in the back with the other." [Laughter] That's Mark. Do you want to sit, or do you want to stand here while I speak? What would you like to do?

>> Mark Eichorn: No, I have to do one thing before I return to my seat.

>> Jon Leibowitz: Oh, you're doing the -- Oh, go ahead. That's fine.

>> Mark Eichorn: When you start the clip, I'll run the clip for you.

>> Jon Leibowitz: Oh, you're going to do the clip? Okay. So, I'll start.

>> Mark Eichorn: Yeah. But I'm standing unobtrusively back here.

>> Jon Leibowitz: So, let me -- [Laughter] Mark has always had my back. There's no doubt about that. So, anyway, good morning, and I want to welcome all of you, both here in Washington, D.C. and those watching online, to today's workshop on facial recognition technology. And we're going to start with a brief clip from "Minority Report" with Tom Cruise. It was a Steven Spielberg movie from 2002. We need some volume. Hmm. Okay, we're going to try that one more time. It's a brief clip.

>> Mark Eichorn: Okay. It's going to do it.

>> Woman #1: A road diverges in the desert. Lexus. The road you're on, John Anderton, is the one less traveled. Make sure nightfall... [Announcers speaking at once]

>> Man #1: You can move the old-fashioned way. Century 21...

>> Man #2: ...provides gourmet cuisine...

>> Man #3: John Anderton, you could use a Guinness right about now. [Speaking continues]

>> Man #4: Stressed out, John Anderton?

>> Woman #1: Get away, John Anderton. Forget your troubles. Escape from it all.

>> Jon Liebowitz: So, I think we can stipulate that our technology is a little clunkier than facial recognition technology. [Light laughter] So, that is the future DreamWorks imagined for 2054 in 2002, just 9 years ago. Facial recognition technology then was the province of science-fiction writers and futuristic movies. In fact, were we having this conference in 2002, we would probably be holding it in Los Angeles. The audience would be full of science-fiction buffs. Steven Spielberg would give the keynote, and we might actually have the budget to provide you with lunch. [Light laughter] Sorry. It's part of or government austerity program. But here we are in 2011, a full 43 years ahead of schedule, with reports of companies beginning to roll out smart signs and tailored messages based on passerbys general attributes, age, and gender, all gleaned through technology that doesn't appear so far from what we saw briefly in "Minority Report." Although we're not aware of any companies engaged in the exact practice depicted in the "Minority Report" clip, including the ones who pay for product placement in the film -- and as you could see, it was like a little 20-second roll of product placement -- the technology isn't that far off. And we'll discuss current uses that even Steven Spielberg couldn't have dreamed of. For example, we have representatives here today from SceneTap, which produces an app that uses facial detection software to help bar hoppers in Chicago scope out the crowd AT more than 50 different bars and restaurants. Now, I suppose if I were 25, that would be a must-have app. [Laughter] And Facebook launched a new facial recognition technology to help you troll through vast stores of pictures, to tag people in the photos you post, or to help others tag you in the photos that they post. And I was driving my daughter to school this morning and I asked her about this and she said she's using this technology all the time. And she says most of the time they get the right photos and match them up, not always, but it enhances convenience. So, these sorts of technologies have already taken hold, certainly in law enforcement and the military. In that area, they are as controversial as they are interesting. But the FTC approaches facial recognition technology from a slightly different perspective, and so will today's workshop. We will focus on the commercial use -- that is, on the possibilities that these technologies open up for consumers, as well as their potential threats to privacy. Most of you know this, but the mission of the FTC, of course, is to protect the

nation's consumers as we navigate the marketplace and to protect competition as it helps to shape the economy. In that role, we walk a line between encouraging innovative technologies that are reshaping our society and protecting consumer's right to privacy, a right as old as our Constitution and certainly as old as the FTC. One of its founders was Justice Brandeis, who wrote about, in Olstead, the right to be let alone, the most sacred of rights, the most important of rights, and the right most valued by civilized men. Is that right? Is that correct?

>> Mark Eichorn: That's good.

>> Jon Leibowitz: Thank you, Mark. Mark found that quote for me once. We do that through both policy and enforcement. Last December, we issued a draft privacy report re-imagining how we can protect consumers in an increasingly digital and mobile age. We expect to put out a final version of that report by the end of -- in the next month or so, I would say. The preliminary report recommended that companies build privacy into their systems as a design phase -- It's what we call privacy by design -- that they simplify the ability of consumers to exercise choices, and that they improve transparency of information practices. And as one example of simplifying choice, the report recommended implementation of Do Not Track, a system we envision empowering consumers -- and not the government -- to take control of the data companies collect about us online. In the enforcement arena, for the three of you in the room who haven't heard about our recent settlement with Facebook... [Laughter] ...our order requires the company to honor its privacy promises and implement a comprehensive privacy program. We've reached similar settlements with Google and other companies in the past. To be sure, the FTC will vigorously enforce the law if we see a violation in the face, as it were, of the -- Yeah, that was a bad pun -- of the new technologies we will examine this morning. But today's discussions aren't focused on enforcement. Really, they're focused on policy. How can we get the most from facial recognition technology without compromising people's privacy rights? Much of what we'll talk about is exciting stuff, really worthy of a DreamWorks screen treatment. For example, this kind of facial recognition technology could be used to find a missing child by, say, comparing a photo taken in a store to one in a missing children's database. But we also have to -- and I think we have to do this right from the start -- sort of acknowledge and address the fact that these new technologies have an enormous potential, really, to run right over consumers if they're not used properly, forcing us to

reveal more than we want to or even more than we know we have to reveal. We have to confront openly the real possibility that these technologies, if not now, then soon, may be able to put a name with the face, so to speak -- again, I apologize for the pun -- and have an impact on our careers, our credit, our health, and families. But fortunately, we are incredibly fortunate, lucky, to have a really terrific group of people here today to begin the discussion. We'll hear from technologists at NIST, the National Institute of Standards, and Carnegie Mellon University, from representatives of small startups like SceneTap and large companies like Facebook, Google, and Intel, from consumer advocates and privacy professionals, out of places like EPIC and the Center for Democracy and Technology. Representatives from the Canadian government are going to discuss their experience with the technology, and my fellow Commissioner, Julie Brill, who hails from almost Canada, will deliver remarks, as well. So, there's a scene near the end of "Minority Report" in which Tom Cruise's character says -- and I quote -- You know your own future, which means you can change it if you want to." Today and going forward, our job is to examine and talk honestly about the future of facial recognition technology and to work together to ensure that it benefits consumers and the marketplace, while respecting all of our right to privacy, the right to privacy for all of us, and the right to control our own personal story. So, with that, I will cease speaking. Do we have a break, or do you want to go right to the panel?

>> Mark Eichorn: We'll go to the panel.

>> Jon Leibowitz: We're going to go right to the panel, and here is Mr. Mark Eichorn to take you there. Thank you. [Applause]

>> Mark Eichorn: Dr. Phillips, Dr. Gross.

>> Dr. Ralph Gross: Do you want me to come right up or --

>> Mark Eichorn: Oh, hold on.

>> Man: Okay.

>> Mark Eichorn: Thank you, Chairman Leibowitz. We'll jump right in to our first panel. We wanted to set the stage for today's discussions by talking about the basics of the technology and the limitations of the current technology and how the systems work. And we're very pleased to have two experts, both thought leaders in this space, here today to talk about the technology. Each will give a presentation, and then, time permitting, we'll take questions. So, for questions, there were some cards in your folder that, if you have a question, you can write it down and raise your hand and FTC staff will pick it up. If you're watching on the Webcast, you can E-mail a question to facefacts@ftc.gov. And we will try to take questions to the extent we have time for them. Let me introduce the two panelists. First, Dr. Ralph Gross is postdoctoral fellow at Carnegie Mellon University, as well as the founding partner and chief scientist of BluPanda and Disruptive Robotics, which we can all agree is a great band name. [Laughter] Dr. Gross' research interests are in the area of computer vision, biometrics, and data privacy. He holds a PhD. and two master's degrees. Dr. Jonathon Phillips, to the far side, is an electronic engineer at the National Institute of Standards and Technology and one of the leading technologists in the fields of computer vision, biometrics, and face recognition. He runs challenge problems and evaluations to advance biometric technology and was awarded the Department of Commerce Gold Medal for his work in this area. He's a fellow of The Institute of Electrical and Electronics Engineers, an international advanced-robotics program. Dr. Gross, let's begin with you.

>> Dr. Ralph Gross: Do I get the clicker? Okay. All right. Thanks, Mark. Let me start by also thanking the FTC for bringing us together here for what looks to be a very interesting workshop. Why is face recognition difficult? It such a hard problem because the appearance of a face changes dramatically with the conditions under which an image is recorded. In order to demonstrate this, here's an example. I imagine that most of you are not going to be able to recognize the person in this image here. You are probably going to recognize him, though, in this image. It's the actor Robin Williams. If you look at the example images below, you see how his face looks very different across conditions such as pose, as illumination, as expression, and as well over time. What's also interesting here, though, to note is that with the exception of his bearded appearance, you probably recognized him in all of the other images, giving you an indication of how well human face recognition works. Face recognition is a very important task for humans, very important social task. There's evidence that there are actual dedicated processes in the brain that

just deal with face recognition, that aren't used for any other object recognition tasks. And as a consequence, humans are very good at recognizing familiar faces. So, if you look at the images that are shown in the middle here, you probably all recognize them, even though they're incredibly low-resolution images. However, if you look at the image pair below and you ask yourself, "Is that an image of the same person?" it's probably a harder task because humans are not very good at recognizing unfamiliar faces. This has been shown experimentally even for people with extensive forensic background. Now let's look at how automatic face recognition developed. By now, we're looking back at more than 40 years of research that started in the 1960s with a system built by Bledsoe and colleagues, which wasn't an automatic system yet. In that system, a human operator had to enter facial-feature point locations for the system to work. The very first automatic system was described by Takeo Kanade in his PhD thesis in 1973. And he reported results on a grand total of 20 subjects. So, we've come a long way since then, as evidenced in thousands and thousands of academic papers, across a range of academic disciplines. When we are looking at face recognition, we are oftentimes looking at one of two application scenarios that I indicate here -- identification and verification. In the identification case, you're given an image of an unknown subject, and the task of the algorithm is to compare it to a set of images of known subjects, oftentimes referred to as "gallery." And the algorithm then outputs a rank-ordered list of matches, hopefully with the accurate match coming out at the top. In the verification case, a subject claims an identity -- in this case, Robin Williams -- and the input image is compared to historic representation of that person and the algorithm gives an "accept" or "reject" response to that claim, depending on if it finds the match to be close enough. Now let's go to the steps that a face recognition system typically goes through in order to process a face. There are four distinct steps -- face detection, normalization, feature extraction, and matching. The task of face detection is quite obvious -- find all the faces that are in the image. This is, again, a difficult problem because faces are complicated objects. What makes it easy, though, is the fact that faces look very similar to each other. So, we all have two eyes, a nose, a mouth, and they're all in approximately the same configuration. It also helps that faces are very different from non-face objects, such as trees or cars or anything else in the environment. Now, in order to detect a face anywhere in the image, algorithms typically look at every possible subwindow in the image by scanning across the image that I indicate here. And as you can see indicated on the right here, this quickly becomes computationally challenging because you're looking at tens of thousands, potentially hundreds of thousands, of individual subwindows,

most of which will not contain a face. So, it's difficult to do that quickly. Now, the first algorithm that really did that in real time is the Viola-Jones detector that was introduced in 2001. And they did, among other things, something very clever, which is that each individual subwindow has to pass through a range of simple filters in order to be recognized as a face. So, the first filter here looks at a little black region on top, a little dark region on top of a light region, which is the eyes above the cheeks. The second one looks at the eyes across or next to the nose bridge, and others follow. And only if an individual subwindow passes through all of these filters, it's detected as a face. So, now, if you look at non-face subwindows that I show here, you see that they very quickly get thrown out. So, only minimal processing is being expended on them. And as a consequence, the detector runs confidently in real time. Here I show a few example detections. You see most of the faces are found. It's actually exactly 90% of them, with three misses and one false positive. That is in the soccer ball right here. Once we have detected the face in the image, the next task is to normalize it. Here the idea is to throw away all unwanted variation. So, what is typically being done is to re-scale the face to a standard size, to rotate the head in case it was tilted, and to throw out any elimination or other imaging variations that aren't helpful for the recognition task. Once the face has been normalized, we extract features from it. Now, the goal of feature extraction is to take what, in this example here, is 12,000 pixel values, which is the size of the face region, and compute a mathematical representation of it that's typically a bit more compact. And there are different ways of doing that. In what's known to be as the holistic or appearance-based approach, we use the entire image -- the entire face region, I should say -- in order to compute a feature representation. In the example here, an algorithm that's called principle component analysis is used, which encodes the image as a linear combination of the images on the right-hand side here. And the features are then the coefficients of that representation. Different approaches called feature-based, where we're looking at only specific areas in the face, typically around facial landmarks such as the eyes, the nose, the mouth, and compute a representation only around there. And there are hybrid approaches that look at image patches, again, usually around salient face points, while excluding textureless regions that are typically not helpful for recognition. So, once we get to this point, we have a compact mathematical representation of the face, and then we'd like to match it to something else. Now, if we're looking at the identification scenario, what you're given is a set of feature representations of the faces in your gallery, and for a new input face, you want to determine, "Well, what's the best match for that?" Oftentimes algorithms do that simply by computing a

distance. So, we're adding mathematical representations here. We can compute a distance. And there are different options of how you compute that distance. There are other ways, too. Hopefully you come up with the right result here. There are advanced models too that you can use, especially if you have multiple images per subject that usually help you in doing a better job in recognizing. So, now these are all the steps that face recognition systems typically go to -- face detection, normalization, feature extraction, and matching. Now let me conclude here by pointing out two recent trends in the field. One has to do with data. For most of its history, face recognition was lacking data because it's very cost- and time-intensive to recruit subjects to come into a lab and sit down and possibly come multiple times to be recorded. And so, as a consequence, there were only a few databases available of reasonable size. Now, this has all changed in the last 5 to 10 years, with the increased use of digital photography, as well as the users of online social networks to share images. One of the recent examples that we're probably going to hear about more today is from face.com, who say that they have indexed 31 billion -- yes, that's "billion" with a "b" -- 31 billion faces in their processing, which I'm sure helps them tremendously in building very accurate face models. The second trend has to do with pose. As this video here indicates, faces are three-dimensional objects. So they look very differently, depending if I look at the face from a profile view or from a frontal view, which makes it very hard to match faces across pose. On the other hand, in real-world images, you oftentimes find pose variations since people don't always just look directly at the camera. So, what's been interesting in the field is that over the last few years, a number of systems have been proposed and tested that are able to compute a full 3-D model from a single input image completely automatically. And what that helps you to do is to essentially re-render a face in any pose you want. I showed in the example here where you take a frontal-input image, you build a 3-D model, and you compute a profile view of it, Which helps then tremendously in recognizing faces across pose. And with that, I'd like to conclude. [Applause]

>> Mark Eichorn: Dr. Phillips?

>> Dr. Jonathon Phillips: Thank you, and thank you for inviting me here and giving the opportunity to speak about the basics of facial recognition. So, I'm going to start off, and a lot of the first couple slides echo the beginning of Ralph's talk, but I think it also emphasizes the importance of some of the points that he made. I should say, when I go through here, I do ask

questions to try to keep the audience awake and keep them participating. So, you'll see the next slide. Hopefully the format comes through. And please tell me who this person is. You have half a second. Oops. There goes my slide. The next slide -- [Laughter] Who is this? I'm sure nobody in the audience has trouble recognizing this person. Of course, this is President Obama, and this is a famous person or familiar person. We are very good at recognizing people we know well. But a lot of the work we have been doing or I've been participating in over close to 20 years has been unfamiliar. So, I'm going to show you a couple more quizzes. Is this the same person or a different person? Are these two images of the same person or a different person? Actually, if I remember correctly, these are different people. And the trouble is you're looking at the face. Look at the shoulders and the body. And by the way, I learned that trick from grade-schoolers during a career day. We'll do one more. Same or different? These are the same. And these images I chose here were selected to be very hard for algorithms, and they also turn out to be very hard for humans. So, this shows that when you start moving from beyond familiar to people you don't know, the problems suddenly become significantly more different or significantly harder in different qualities. So, the first thing is just a few performance statistics. This is the verification task, the one we showed before where, "Are these the same people or different?" It's identification, which is the first person when I asked, who is in the picture of President Obama. So, the first one is false accepts. There are two types of errors. One is -- you saw in the first one -- you think different people are the same. This is the false accept. The case I have here occurs when a system accepts an invalid identity claim. For example, if I were to claim to be Ralph Gross and the system says "yes," that's a false accept. Now, there's the false reject. If I go up to the ATM and say I'm Jonathon Phillips and it says "No, you're not," that's a false reject. I should emphasize there's a tradeoff between these two. So, I can let nobody into my system, in which case I'll have no false accepts, but I got 100% false rejects. The other end of extreme is, I can let everybody in. I'll get no false rejects, but my false-accept rate would be 100%. And there's a tradeoff between these two as you set the system parameters. So, I've been working in face. I first started the FERET program in 1993. And we were addressing this particular problem -- giving two images - they're nice frontal images taken in a mobile studio, a mug shot -- and we want to compare these. And this question, "Are these two images the same person?" I hope everybody says that they are. And the next thing I'm gonna show is that -- So, what I've done is -- At NIST, we measure performance or we provide -- Or the service that I've done is encourage the development of technology. And particularly, we

have concentrated on this problem. I'll show you the next slide, but there's more to the story after this. So, between 1993 and 2010 is a remarkable story of progress in the technology. When I first started in 1993, I had a false-accept rate of 1 in 1,000. So, that means -- What happens if we accept the rate -- so, we only accept 1 in 1,000 false claims. What's the reject rate? Well, when I started, it was 79% false-reject rate. And after the first successful program, we got it down to 54%. And this is the FERET database, which is in blue. Next, we moved in 2002 with the Face Recognition Vendor Test, 2002. We moved -- Watch in red. This is the Department of State non-immigrant visa database. This had over 134,000 images, if I remember the numbers correctly, 36,000 people. So, this is operational data of performance. And what we see now -- in 2002, it was 20%. Over the last 8 years, it's now gone to .003, the rate, or 1/3 of 1/10 of 1% rate. So, this is a phenomenal decrease. The error rate over this 17-year period was, we half the error rate every 2 years. So, for many applications where you're dealing with mug shots, phenomenal performance. The other number we did in the last evaluation, 2010, we had a database of 1.6 million mug shots provided to us by the FBI. And we did rank one retrievals. So, we'd took an image of somebody in it, as was the first rank one retrieval, and we get success rate 93% of the time. So, as I usually talk in front of researchers, I go, "Rest assured -- there's still research problems out there," which is now what I'm going to discuss, is start moving beyond the controlled images you see in the previous one -- these examples of me because we can't show some of the real operational data. So, the question is, let's move to other imaging conditions and sort of have a green tab over there. You've seen green. This is taken in a controlled environment. We have a solid green. We start moving to ambient light or a room like this or outdoors. It gets a bit tough. In the same condition, we start moving rotation some, and yellow, more difficult. And actually, if you start moving outside with people in shadows --you see at the end -- this is a very though, tough problem at the leading edge of research. And the other thing I'll go to is identify some of the challenges, and a lot depends on what we call inequality of images. We created three performance levels -- good, bad, and ugly. These were taken with a D70 Nikon camera. This is a digital SLR. It takes very nice quality images. These are indoors and outdoors, frontal faces, and taken within a year. So, here are examples. The key thing to think here is, it comes in pairs, what comparing to. So, these are six images of the same person. I think over here. I'll try here. So, this is a pair that were in a good set. I can't go all the way over there. Sorry. The middle one -- these are challenging. Same person. And these are very challenging, moving outside. So, you start to see some of the difficulties in moving outside the areas of advance. Here's

another same person. And a different person. There again, these are six images of the same person. So, now we have performance. Once again, this time, this is the verification rate, successful recognition for good. We're 98%. These are numbers from -- at least, the top ones -- from 2006. The numbers of the top ones have probably gone up. This is very good performance. You see many uses for it. The bad -- down around 80%. Do we have any takers for what the ugly is? I know there's some people that have seen this presentation before. Somebody who hasn't, can you give me a guess? So, this is 15%. So, this shows the outer limits of researching the problem. This is a tough problem. We've encouraged them to do research. So, this sort of shows the full range of images you can expect, depending on the quality in what you put up. So, the next one here are examples of point-and-shoot. So, sometimes you need to look at a lot of these. These are a database which we're putting together to put a challenge problem out at NIST, were collected. They were all the same person. Professor Kevin Bowyer at Notre Dame -- you see now the full range of performance, you can see. So, this is kind of showing the question that came up when you put these together is -- You know, a lot of the work we've done is pushing us head with the mug-shot type things. So, the question is, as you start moving to social media, and the question is, "What type of images you get with point-and-shoot?" So, I sort of give a few examples of what we see out there and where we're going and the difficulties that have been in processing. So, the last one of the questions they addressed was the question of aging. So, I've put this up here for two points. Across there was when I was working the FERET program, early in the database. This was taken in the mid '90s. And below is an image of me taken about 50 years later that I now use as a picture for some -- So, one of the issues that came up is we recently did an evaluation at NIST in the MBE2010, where we had mug shots taken up to a period about 9 to 10 years. And what we found is the best algorithms were invariant in performance. They did not change up to about eight years. So, this shows that, up to about eight years, we have quantitative performance. Unfortunately, we don't have any solid performance on numbers beyond that eight years. In other words, generally these, you know, 10 years, 20 years can be very tough, as we change a lot. But the other issue I'd like to say that I think is more significant for a lot of the aging, at least within the 10- to 20- years period, is I think changes in performance will be dominated by changes in illumination, change in focus, change in camera and pose. So, in many cases, the question about aging would be not concentrating so much on "What's the pure effects of aging?" but question is that the changes in images are more dominated by other photometric concerns that do come up. So,

I'd like to, I guess, move to question time, but I think I would just like to summarize the two key points, is that there are two fundamental classes of face recognition. One is recognizing familiar faces. We're all very, very good at that, and we see new faces or make decisions. Sometimes we interpolate our decisions or extrapolate our decisions based on our ability to recognize familiar faces versus unfamiliar, which is a different problem, which a lot of algorithm-development community effort has worked on. The other is, we have mug shots or controlled frontal illuminations. There's been fantastic progress in the last 17 years, as shown by missed evaluations. But as we now move and solve that problem, we're now moving upon to the other cases. What happens if we do point-and-shoot? What happens with video and a lot more constrained? In many ways, this is the cutting-edge of research we're just moving into, and it's probably -- probably what I've seen more models of the southeast social-media concerns being addressed. And I think this is now an active area of research, and time will tell over the next five years if we see similar limits in performance or maybe the different capabilities. Thank you. [Applause]

>> Mark Eichorn: Thank you both very much. Let me, I guess, start off the questions with one sort of taking off on the "Minority Report" theme -- and you both addressed this in different ways -- but with the increase in images that you talked about and the, you know, vast increase, the acceleration of the accuracy rate in the last years of study -- Is that "Minority Report" kind of situation feasible now or feasible in, you know, 10 years, with sort of predictable advances in, you know, the number of images, the increase in computer speeds, the reduced costs of computer storage and developments such as that, as well as the developments in the facial technology itself?

>> Dr. Jonathon Phillips: So, you -- Oh, this is -- Oh. Oh. So, one of the things about "Minority Report," it just knew who that person was moving forward, which is a huge database. I think, you know, these being frontal -- I think for the moving forward, I think that'll be always a challenging problem. I think the limits of where you can push face recognition in unconstrained situations is an open question. In other words, there's some fundamental limits -- that we don't have "X" quality of images, where I mean quality, not -- or maybe quantity over enough viewing conditions. I think it's an open research question of what the limits of recognition are in the unconstrained environment.

>> Mark Eichorn: Ralph, do you have any comment?

>> Dr. Ralph Gross: Yep. We're certainly see that we're moving towards that direction. There are already commercial products being installed that equip billboards with face processing capabilities. So, they don't do recognition, but they try to do things such as gender classification or age classification in order to tailor the advertisement that's being shown to that particular demographic. So, we're already seeing that the environment is being equipped for that. Now, you know, I agree with Jonathon that we're certainly not there yet to run that in real time across a database of 300 million people. But as we've just heard, the performance increases have been quite dramatic over the last 10, 15 years. So, if that continues in the same direction, then, you know, this certainly might become possible.

>> Mark Eichorn: Just a plug for the next panel -- we'll be discussing a lot of those facial detection technologies in the next panel. Are we at the point where computers are arguably better at recognizing faces than humans are in some situations?

>> Dr. Jonathon Phillips: So, over the last, I guess, six, eight years, I've been collaborating with Professor Alice O'Toole, comparing human-machine performance. One of the caveats is, these are performance of people not specifically trained to recognize faces. And when you deal with frontal face recognition, for example, with the good, bad and ugly and other evaluations that we've done, machines are better than humans at recognizing faces, in the frontal face recognition across changes in illumination. But overall, if you look at the entire face recognition capability, humans are still the most robust face recognition algorithms out there.

>> Mark Eichorn: So how do certain factors such as -- You know, lots of us have these digital cameras now that, you know, may be five megapixels or seven megapixels. Are we at a point where we're sort of at enough density where any further improvements wouldn't significantly change the accuracy of these systems? Or is there still room where -- like, if we have denser pictures 10 years from now, that will have an impact?

>> Dr. Jonathon Phillips: I'm trying to make sure they saw it. So, the interesting thing is the -- as we move to point-and-shoot, many times we have very 10-plus megapixels in the camera versus the

older ones I've shown taken with a digital single-lens reflex are maybe six megapixels. And the conclusion I've come to after looking at these is, it's not the number of megapixels you have. It's the quality of the camera and the lens, the optics, and the ability to F-stop down itself. So, the issue is not how many more megapixels you can stick on your smartphone. It's, can you stick a better lens and focusing mechanisms on the smartphone that dictate the performance over a wider range of the quality of images for recognition.

>> Mark Eichorn: We have a question from the audience, and the question is to Dr. Phillips. Has NIST issued any standards on performance quality or integrity of facial recognition technologies?

>> Dr. Jonathon Phillips: So, one of the research projects we're starting out with is to measure quality -- to start measuring quality of images for recognition. This is not a solved problem. We have, for example -- I should say we're in the process of formulating a challenge problem to be able to measure and assess what are effective quality measures for face recognition for matching.

>> Mark Eichorn: So, another question is, as we get more and more images of each of us and images that are associated with us, does it then become easier to identify an unknown person when you're comparing to sort of a known data set? If you have 100 pictures of me that are known to be pictures of me, is it easier to identify that 101st picture of me?

>> Dr. Ralph Gross: Yeah, if you're doing it right, it becomes easier, yes. So, as we've shown, the face varies a lot with different conditions. If I have very fine images of you under a lot of conditions, I can build a better model of how your face looks like and have a higher chance of having seen you in a particular condition when a new image comes in. So, if I have a single image of you, it might be bad lighting. It might be bad pose, so the chances of using that is not as good as if I have 100 ideally varied images that show you from different points of view, under different illuminations, and then the chance of having a positive identification is much higher.

>> Mark Eichorn: So, what are the implications of that? If you have video of me walking down a hall, then you would presumably have some pictures of me in shadow and some pictures in good

lighting and so forth. So, does that mean that it's easier to use facial recognition on video? Or is it easier with a perfect studio still photo or...?

>> Dr. Jonathon Phillips: So, I think -- I don't know -- The question is having -- whether -- Probably a perfect ideal frontal image would be better than a video, but the other thing to remember is that probably the differences for performances would be not -- having one video would be nice, but probably what's more important is having multiple videos or multiple images of you under different conditions in different times, is probably more important than -- would contribute to improving the performance than just a single video itself.

>> Mark Eichorn: Here's another question from the audience. The video clip from "Minority Report" actually portrayed the use of retinal scanning. So, it was not technically facial recognition, but the iris was being scanned. So, what is the difference between retinal scanning and facial recognition?

>> Dr. Jonathon Phillips: So, anyway, I'd say retinal scanning and iris scanning are two different things. So, iris scanning is if you go look -- If you look in a mirror, you see all the different patterns. It's generally believed and the evidence is showing that each one of us has unique iris patterns itself. The excess of iris was started by John Daugman at the University of Cambridge, is that, there again, you take the iris images to get the very high performance under very controlled conditions. And we have finished evaluations at NIST -- I believe it's called IREX III -- which should be posted, showing performance on very large databases. But these are all very controlled. But as with face, as soon as you start moving away from controlled situations, at a distance, the performance does drop. And there are challenges with iris about actually imaging the iris itself at the distances that you see in "Minority Report."

>> Mark Eichorn: So, to follow up, would an individual more or less have to participate in that kind of scanning process? I mean, I would have to either be within a foot of a scanning device or something like that?

>> Jonathon Phillip: So, the ones that are cooperative, you actually go up and interact, like a foot away from the scanning to get a high quality. I would say acquiring irises at a distance is currently a research topic.

>> Mark Eichorn: And you said that retinal scanning was a different type of process?

>> Dr. Jonathon Phillips: Yeah. Basically, I think, if I remember correctly, retinal looks at the blood vessels in the back of the eye, and it's much more intrusive than retinal scanning. I don't know of any serious applications with it.

>> Mark Eichorn: Okay. Could you talk about the differences between facial detection and facial recognition? You discussed it, Dr. Gross.

>> Dr. Ralph Gross: In terms of performance?

>> Mark Eichorn: Yes.

>> Dr. Ralph Gross: Yeah, generally, face detection performs better. You see it quite frequently in use in your consumer-grade cameras, that you have a little green box hovering over faces. And across a large set of conditions, we've seen algorithms that perform in excess of 90% accuracy. So, I think, considering that you're comparing two different problems, it's always a bit of an apples-and-oranges issue, but generally it performs very well.

>> Mark Eichorn: What about the possibility of detecting emotional states, like whether somebody is smiling or frowning or...?

>> Dr. Ralph Gross: There are algorithms out there. And it's a reasonably active research area that looks at facial expression recognition. I don't know if we can make any statements as to how accurate that works. And certainly one of the issues is that, you know, what's the application, really, of it? And the papers, as always, been described as, your computer wants to be more helpful and recognize your state of emotion. But I don't know that I buy that.

>> Mark Eichorn: And what kind of criteria -- We've heard of age and gender being used or detected, some type of categorization based on that being made. Are you aware of any other types of sort of identifications that are made by facial detection?

>> Dr. Ralph Gross: I'd say these are the most popular one, gender and age. You know, ethnicity is certainly one that I've seen, although not as frequently. So, yeah.

>> Mark Eichorn: Okay. Well, with that, we'll wrap up the first panel, but I really appreciate both of you coming today, and thank you for being here. [Applause]

>> Mark Eichorn: And we will have a 5-minute break. [Indistinct conversations]