MITOCW | mitmas_531f09_lec05_3.mp4

[SIDE CONVERSATIONS]

MICAH KIMO JOHNSON: OK. Let's get started, again.

[SIDE CONVERSATIONS]

I'm going to tell you guys about a recent project that I did, which was presented at CVPR this summer. It was a paper there and a demo, and I also was part of Siggraph Emerging Technologies, and it was pretty well received there too.

AUDIENCE: [INAUDIBLE]

MICAH KIMO JOHNSON: We're starting.

AUDIENCE: OK.

MICAH KIMO JOHNSON: The main idea is a technique for capturing geometry. I have here, this is our little pocket demo, which I'll pass around, but it's a really simple idea. This is a piece of gel with a painted skin, and if you press something into it, like your finger, it looks like it's been painted. So I'll pass this around. You just take turns pressing your finger into it, looking at it.

It's pretty simple and pretty cool, and what I'll tell you about in this talk is how we can actually take advantage of this simple device to measure surface geometry. Now this is been a pretty long standing problem in vision, how can you capture geometry? How can you get more than a 2D image? For example, how do you get 3D shape?

And I have some references here, probably not so interesting for this class. Basically, for these different techniques, they are thinking more generally. How do I capture geometry of somewhat general objects? And I'll show you.

This has some very specific properties. It's only going to be able to capture surface texture and surface geometry, but we still think it's pretty cool. For example, it can capture geometries of objects like this. We have an Oreo cookie here, so the surface of it has some interesting detail.

This is a glass little trinket. It's like a pin. And a lot of existing methods that either use lasers or use structured light have problems with glass, so that's a difficult type of object to scan. This is a metal drawer pole. It's just a small piece of metal, but techniques that use cameras and structure light also have problems with metal because of the reflectance function.

And we can even scan the surface geometry of something as fine as a $20 bill, so I'll show a lot of different results. Here's the main idea, again. You have some of this clear gel, called an elastomer, thermoplastic elastomer. We put some paint on top, when you press something in and view it from the other side, it looks like it's been painted. And when you get the pocket version, you can see that yourself.

So we can use standard computer vision techniques to reconstruct the surface geometry, and I'll tell you about that in the rest of the talk. But first, what is this material? Well, it can be any kind of clear gel, so we use thermoplastic elastomer mostly, which is a common rubber that people haven't heard of, but it's the same kind of rubber that's in your shoes or in a lot of different products you buy. But you can also use silicones and other rubbers. It's just important that it's clear.

And we bake it in the oven. We melt it down into whatever shape we want, and typically, we want it to be flat on top. But you could make it into the shape of your face because you can use this stuff to make molds, so just some interesting things we've done making it to other types of shapes. But typically for this purpose, we want it flat.

We engineer a specific paint to put on top, and you can see these are four different sensors, four different paints. And the properties of the paint kind of change the types of geometries you can measure. I'll maybe talk about that a little more.

But what we do is we put this into a photometric stereo setup. There's some gel on the top. This is a box that I made. And you've got a camera beneath it. And then inside, there's a Tupperware bowl with three lights, red, green, and blue from different directions.

Let me show you actually a video. So as part of the Siggraph presentation, I actually made a few videos, and let me play them because you can see it in action and see how it works. This is a real time capture, so this is the top of the box. You're zoomed in. You're looking at my finger pressing it to it. On the right is what the camera sees.

This is looking at the bottom of the gel through glass, underneath the bottom. We have three colored lights around it, red, green, and blue, at different directions. And if you imagine what each of the individual color channels looks like in this image, if you just look at the red channel, it looks like a shaded image with light coming from this direction. If you just look at the green channel, the shaded image with light coming in that direction, same for the blue channel, shaded image with light coming in this direction.

So by using three different colored lights and an RGB camera, we can have three separate shaded images. It's gray-scale images, but it's three different light directions. And it turns out that that is enough to constrain the surface normal on the shape of the top of the objects.

I can talk a little bit about the math after this, but just to give you an appreciation for what the demo is like-- so this is what the camera is seeing as I move my finger around, but we can do a reconstruction of the geometry in real time. So I'm spinning the viewpoint, and this is all captured geometry and another camera feed just to see the interaction. This is what looks like when you brush your teeth. And this is the reconstructed geometry of the toothbrush and its bristles.

And so this is the demo that we did at Siggraph, and also I just did it at [? ITCP ?] last week. And you can see the Oreo, interesting properties. It's high resolution. It's interactive, so you can get dynamic 3D information.

Now this is after the Oreo. I noticed there were crumbs on the sensor, so you can see the geometry of the crumbs as they roll underneath my finger. And this is Ted Adelson's pulse. He's put his wrist down on the sensor, and he's reconstructed all of his skin texture and also the pulse that's moving.

And finally just for fun, the surface of the bill, and this is all done in real time. I have to press it with a flat plate, and you can see-- maybe not those of you who are looking at an angle, but if you're looking at it straight on, you can actually see the raised printing. So it's got a lot of different properties. I said, real time, high resolution, but it's not a traditional 3D scanner because you're not going to be able to scan your face or something with a lot of depth, can't imagine pressing your face down to the gel to get full 3D. We don't claim it's a 3D scanner anyway. It's really [? 2.5 ?] [INAUDIBLE] a type map in real time.

So let me go back to the presentation. How does it work? Well, why do you need three different light directions? Why don't you just have one light and try to reconstruct the shape with one light?

Well, that's called shape from shading. And in computer vision that's considered one of the traditional hard problems. So we don't want to have to attack a hard problem, especially if we want this to happen in real time.

We use photometric stereo, which is actually a very old technique too, first proposed by Woodham in 1980. And it says, if you have three lights-- and this is a Lambertian sphere lit from three different directions. So you can see the lights are in each of the color channels red, green, and blue. They provide constraints on the surface [? level. ?] We'll talk more about that.

Here's the three shaded views, just so you can see. The Oreo's pressed into the sensor. This is an actual picture that I took with it pressed in the sensor, red channel, the green channel, and the blue channel. And you can see the lights coming from different directions.

AUDIENCE: Why don't you have broad sources? They're not--

MICAH KIMO JOHNSON: Yeah, they're actually sort of little area light sources, and we found that we get the best looking results in terms of detail by using specular paints, which have a lot of contrast, with area light sources. So there's a lot of things you can fiddle with. You could go to a point-light sources with a Lambertian or a very diffuse paint, and that's better at reconstructing a lot of depth. But to get fingerprints and details on the surface of a bill, we found that it works better to have glossy paint and broad lights. But, yeah, all sorts of things you can fiddle.

Actually, let me just go into a little bit of detail why you need three lights just because I think it's interesting. So again, as I said, if you'd only had one light, you'd have to solve shape from shading problem. Why is this hard? So here's a shaded sphere with a single light direction. This is a Lambertian, so it's a diffuse sphere.

And let's say you know that the intensity value of a certain pixel is 0.9. Well, here I find a [? iso ?] [? platonic ?] sphere, so [? 0.99 ?] here. But in fact, the orientation that corresponds to 0.9 [INAUDIBLE] because all of these points having a [? normal ?] that makes the same angle of a light direction. And that's how a sphere shaded in a diffuse kit.

So just knowing the shading at a particular pixel that is 0.9, it doesn't tell you anything about the orientation of the surface underneath. But by having three light directions, if it's 0.9 in green, 0.8 in red, and 0.7 in blue, or something, you can pinpoint-- we have the intersection of these, these types of [? ice ?] [? coats. ?] And you can pinpoint the location on these [INAUDIBLE] three lights.

So it's really pretty much that simple, where you look at the colors in each channel, and you build a lookup table through calibration. So we calibrate by pressing into it this grid of spheres. We know the size, and we know the geometry because it's a sphere. It's very simple.

You press it in, and it looks like this. And already you can see that color denotes orientation. If you see a pure green pixel, it means the orientation is something like this. If you see a pure red pixel, it means your orientation is down like that, and blue is over like this. And then if you see a mixture of colors, it means it's some mixture in between those orientations.

So you can build a mapping from color to surface normal, so it's a 3D lookup table. We have red, green, and blue indices, and each of them you can get the gradient or the surface normal. So again, if we see color, 10 in red, 200 in green, and 3 in blue, we go to the lookup table, and we get out that gradient, something like that. So you get a different color, you get a different gradient.

So then we solved Poisson equation to reconstruct the surface from its gradients. Ramesh mentioned that before. It's fairly common and used in a lot of different problems. But here are some results.

So these are the results from the [? CBC ?] paper. This would be really challenging to use laser scanning or traditional photometric stereo. But when you press it in, the skin of the gel potentially paints [? sideways. ?] Everything you press into it looks like it's been painting. So we can reconstruct the geometry, and you get-- this is head-on view and a view from a different angle, and it's very direct and straightforward.

So again, here are these objects that I showed before. And you can see we've captured a lot of the details. Now you don't get color information because you're essentially painting it when you press it into the gel. But in a lot of cases for graphics, just having bump maps, having the texture is useful.

And finally, the 20 [? probably-- ?] so for the [? CVCR ?] paper, I did a 20. I stepped it up a bit for the video and did a hundred, but they both work pretty well. So that's the height map that we've measured. And now you can render, and you can change the viewpoint. And you can see that we've captured quite a bit of detail.

And that's just the security strip [INAUDIBLE].

AUDIENCE: Going to get into counterfeiting then?

MICAH KIMO JOHNSON: Yeah. yeah. You can actually sending this to a 3D printer. So recently, we showed this, like I said at Siggraph Emerging Technologies, and we took sort of a graphics slant showing how you can measure surfaces. But this could be used for a lot more. So we just made a second video, and I talk a lot in this one, but let me just fast forward to the interesting ideas.

OK. So in this video, we're going to look at how can we change the properties of the gel to measure different things, perhaps beyond the limits of the original system. So how can we get even higher spatial resolution? How can we change the sensitivity of the system? And can it be made into a smaller device?

Because everyone asked us when they saw this big box that I had, well, you know, they said, that seems really bulky and cumbersome. Does it have to be that big? So we've made a smaller version, but let me step through some of these results.

So spatial resolution, this is an oregano leaf pressed into a very thin piece of this sensor. So it looks like the oregano leaf has been painted. And you can see a lot of detail with your naked eye like this, but then you can put it under a microscope too, and zoom way in. And you can see a lot of these structures, and I don't know the technical terms for them.

But you've got these little pieces here, these hairs. We estimate that, I believe, some of these structures are 100 microns. And then at the tip of the hairs, that's perhaps on the order of just a few microns. So you can really see a lot of detail just by pressing this gel into it.

AUDIENCE: Wait. Is this the 3D reconstruction or just--

MICAH KIMO JOHNSON: This is not a 3D reconstruction, but the same principle could apply. It's just a different piece of gel, and we just took a picture. If we put the three lights in and calibrated, we could get the 3D geometry out of this. Calibration might be a bit challenging and getting the lights configured, but in principle, this is no different from what we did before.

So then we look at sensitivity. A lot of these are just Ted Adelson's results. He's the one who came up with the idea of the gel and this. And he's been fiddling with these things in his basement. I did the 3D reconstruction stuff.

But I didn't even know he was thinking about all these different ideas, and he sent me these videos. So this video is his car driving over a piece of this sensor. So you think, how did he do this?

Well, first of all, why he did it, so I don't know. But the idea is that-- the gel that a lot of people have played with in our demo is soft to touch. So you touch it, deforms, and you can see your fingerprints and things like this. But you can get these rubbers in a variety of strengths, variety of elasticity.

So he went out and got one that was as hard as a tire itself. And then he bought one of these ramps that you used to change your oil, so you can drive a car up on it. He cut a hole in the top, put some of this gel down, put a mirror underneath that had a camera, and then drove his car over it.

You can see these are some pebbles stuck in the tire. This is the view of the tire. Now this is not reconstructed 3D geometry, but again, it could be that. You could actually measure tread wear or something like that. It just shows that the system can handle 2,800 pounds of car on top of it.

But on the other side, this is him pressing, just poking at with a hair. So this is a very soft piece of gel, and he's just touching it with an end of a hair. Or on the other hand, this is some soap bubbles viewed through clear glass. So this container was clear glass and soap bubbles inside. But then he's put sensor material on the other side, and you can see that just the pressure of the soap bubbles is enough to deform the gel.

So on the one hand, we had a car on top of it. And the pressure from the car, to start it, we could get geometry there. On the other extreme, we have soap bubbles, and you can see the geometry, and you can even track how they form and how they change. So there's a wide range of sensitivity to do anything in between.

This is somewhat disturbing. So in talking to-- and he sends it with this in his email, and I open it up, and it plays. It's like [? disturbing. ?]

And talking to technology licensing, they thought, wouldn't this be interesting to do medical imaging? So there's some exams, like prostate exams that are done by a doctor, putting their finger in, feeling, so the touch-based. And then they write down what they feel, but there's no record of that. There's no image that they take, nothing like that.

This is a touch-based sensor that can capture an image of what it touches. So in theory, you can build it small enough, capture the geometry of internal tissues. But rather than make something that we would put internally, Ted just licked it to show internal tissues was being measured.

Let's see. We've got, oh, yes, worms. These are worms on the surface of the sensor. And looking from the top, you can't really tell how the worms move very easily because they're translucent.

You can't see the details, so they're not so obvious what they're doing. But at the same time, he was recording underneath. So this is the view from the sensor. And you can see that as the worm expands and contracts, you see these protrusions here, these structures that stick out and grip.

And I learned they're called [? CT. ?] And perhaps people who are interested in saying worms might be interested in seeing or measuring how they expand and contract and how these little structures come out. So we can just see it visually, much clearer because we've essentially painted the worm.

And finally to address the question of can we make it smaller? Ted made a compact version. And let me explain this one a little bit.

This is a dental camera, so it's somewhat like a toothbrush. But inside here is a camera and a light. It's very small. And so on top he built like a lens and piece of sensor material. You can see that.

It has a little bit of memory in this one. But when he touches it, it's the same principle of what we've done before. It's got some of the sensor material. But now it's in this handheld device, and it could be made even smaller if we just engineered smaller light, smaller cameras. This was just something off the shelf that we bought, a dental camera, so it's pretty small.

And we can show that you just touch it into things, and you can see the texture in geometry. And with three lights, we could do our same trick of photometric stereo and actually measuring the geometry and reconstructing it. But this one is kind of fun to play with.

You can see this might be useful for skin texture. So one of the last remaining challenges of doing realistic human rendering is getting this microstructure of skin, the wrinkles, the pores, all the things like that. And people have built some very expensive domes to try to measure a lot of the geometry, but they can't get down to this level of detail. We can get it locally, but we can't get the low resolution. We can't get the actual structure of the face, so maybe combining these techniques could yield something interesting.

So here's stubble. You can see the stubble actually has quite a bit of geometry, more than you might think, and his hair. That's basically the touch-sensor system. Does anyone have any questions?

AUDIENCE: Yeah, I do have a question. So if you use that little [INAUDIBLE] table to get some information and get the direction in the wrong findings, but you may have several problems separately or from just a different path, and you need to straighten that [INAUDIBLE]

MICAH KIMO JOHNSON: Right, so--

[INTERPOSING VOICES]

AUDIENCE: How would you get that right? You can say, OK. How would you distinguish this [INAUDIBLE]

MICAH KIMO JOHNSON: You're relying on the ability to measure orientation accurately. So we're measuring orientation, the surface normal. And we've verified that for relatively shallow orientations, meaning orientations that basically face the camera we're pretty accurate. As you get steep orientations, we've become less accurate. And we could employ other techniques to have more accuracy there, like maybe more lights, more views. But in general, we're not getting steep normals accurate.

So we reconstruct the depth from the normals, so there might be some bias. Let's say the surface had a dip like this. If you missestimate this normal, you might only reconstruct it like that, so you can get errors like this. But if you're mostly interested in high-resolution detail, this device is pretty good. If you need to get those low-resolution, depth estimates accurate, you might need more views or more lights than what we're using.

AUDIENCE: What is the state-of-the-art in touch senses that you maybe drag over something that aren't camera-based? I mean is there anything in that?

MICAH KIMO JOHNSON: Yeah, I think the state-of-the-art is 10 [? prolimens ?] for [INAUDIBLE] internal reflection. That's something which--

[? RAMESH RASKAR: ?] Defining [INAUDIBLE], NYU?

MICAH KIMO JOHNSON: NYU. But [INAUDIBLE] to it because he was on it.

RAMESH RASKAR: Now you're talking about the--

MICAH KIMO JOHNSON: Oh, there's the Microsoft Surface, but then there's this new--

RAMESH RASKAR: No, no. Ken Karlin has a new force-based--

MICAH KIMO JOHNSON: Oh, it's force based.

RAMESH RASKAR: --multi-touch sensor, which he has shown in--

AUDIENCE: Is it like gel?

RAMESH RASKAR: Yeah, the gel one, and then the [? internal ?] reflection is purely optical.

MICAH KIMO JOHNSON: I was going to ask about that.

AUDIENCE: Is there anything where you directly measure? I can imagine a line of points that you can measure the height that they get to press, and you drag your finger across it. Is there anything like that?

RAMESH RASKAR: How would the height change help you?

AUDIENCE: Like where you're directly measuring the height instead of inferring it based on images.

AUDIENCE: Oh, yeah, there is. The [? Soundscape ?] demo.

MICAH KIMO JOHNSON: I think that that one [? Ken ?] [? Karlin's ?] one does measure pressure, so that you can get pressure variation, directly, I think. [INAUDIBLE] But it's low-resolution. So I haven't seen anything that's high-resolution like ours with fingerprints.

RAMESH RASKAR: They can do a pan, but they cannot do [? reaches. ?]

MICAH KIMO JOHNSON: All right.

RAMESH RASKAR: Any other questions for Kimo?

MICAH KIMO JOHNSON: Yeah.

AUDIENCE: How do you come up with the ideas?

RAMESH RASKAR: So what was the question?

MICAH KIMO JOHNSON: How do we come up with the idea? So I think it's a couple of things. Ted has young children, and he was fascinated with the idea that touch is actually a very important sense to infants. They're touching things to their lips, [INAUDIBLE] lot of sensory organs that they're touching. And touch is relatively unstudied compared to vision and other things.

And one of the reasons is that there isn't a good touch sensor, something that measures touch the way skin does. It deforms when something presses it. And so he thought, well, how can I build a touch sensor, something that deforms like skin. I think at the same time he was having foot problems, so he was just separately buying these gels to make insoles. And then these two things mixed in his head, and he thought, oh, I'll make the gel.

RAMESH RASKAR: Did you pass this around?

MICAH KIMO JOHNSON: I did. Yeah.

RAMESH RASKAR: I was fortunate enough to have one for myself. But if anybody wants to do a project on this, you can collaborate with--

[INTERPOSING VOICES]

MICAH KIMO JOHNSON: Or if you want to actually play with the demo in person, we have it set up in our lab.

RAMESH RASKAR: Or he could be a mentor if you want it to be a final projects in the class.

AUDIENCE: So [INAUDIBLE] have some idea in the space? It can--

MICAH KIMO JOHNSON: We are thinking of what the next generation of this will be, and we're shooting for a Siggraph project based on this.

RAMESH RASKAR: All of our [INAUDIBLE], that's why the classes in the fall. And the final projects are due 4th of December. We have one and half months to watch. Thanks, Kimo.

MICAH KIMO JOHNSON: Yeah.