MITOCW | mitmas_531f09_lec01_2.mp4

RAMESH RASKAR: So let's start with a thermal IR camera. And as you can see, it has very strange properties. Who are we looking at here? That's you?

AUDIENCE: Oh, yeah.

RAMESH RASKAR: [LAUGHS] And then, as it turns on his lighter, it caused an automatic gain mode, I guess. And that's why the image gets very dark. And who's in the picture, anyway?

Oh, that's Seth. Excellent. Thought I could recognize you in thermal IR And if you just put a cold finger on your cheek, you're going to see it.

And take it off. Yeah, you can see the mark. Things that are completely transparent in visible are actually completely opaque.

And the focus is-- yeah. That's completely opaque. Actually, if you just put your glasses down, this-- might as well. You won't see, but others will see.

AUDIENCE: Yeah, I will see.

RAMESH RASKAR: [LAUGHS] So they're completely opaque in--

AUDIENCE: These are actually glass lenses, not plastic ones. There might be some difference.

RAMESH RASKAR: OK.

AUDIENCE: I don't know. But--

RAMESH RASKAR: So it's really looking at about 8,000 nanometer to 12,000 nanometer, which fortunately is also the wavelength in which the human bodies-- the black body radiation of the human body peaks. So you can do some amazing things with this.

Chrysler and BMW are thinking about-- or maybe already put them in automobiles so that you can see things very far away. So one benefit of this, as you can see, is even if you turn off the room light, the intensity will not change. So if somebody must take the-- it's completely independent of the room lighting, because it's not looking at the visible spectrum that's being emitted by these tube lights. But it's a function of just thermal radiation. So at night when you're driving, if there are any animals or even intruders in your backyard, you can detect them with thermal IR

AUDIENCE: There's a lot of reflection on the table.

RAMESH RASKAR: Yes. Yes, because remember, this is a very large wavelength. So that wavelength-- the table's roughness disappears and it becomes highly reflective. Is this-- let's now try to focus on things that count. All the people with glasses look very cool.

[LAUGHTER]

AUDIENCE: What is the lens made out of?

RAMESH RASKAR: So that's a good question, because typically, glass will not transmit thermal IR.

AUDIENCE: Right, [INAUDIBLE], yeah.

RAMESH RASKAR: Yeah. That's a good demonstration. So a lot of these lenses are made up of germanium. And also certain types of plastic can be used for this. But the image quality is not that great. But germanium is very common.

AUDIENCE: Does this have different focal characteristics because it's such long lengths of long wavelengths?

RAMESH RASKAR: It has 8 to 12-- just in terms of the ratios, it's 300 to 400. It's the same as 8 microns to 12 microns. So it should be able to focus in the majority of that band. But because you need a larger lens, the depth of field is pretty narrow.

AUDIENCE: Yeah.

RAMESH RASKAR: So with this, you can build a glass detector. And if I look outside the window, for example, it's completely opaque. Can you wave, the last person?

Yeah. So it's completely opaque, the glass. So you can build a glass detector. You can take a photo with a visible, regular camera.

You can take a photo of this. And if it's opaque in one but transparent and other, then that's glass. So you can do all kinds of interesting things. And again, if Mattias is successful and other companies are successful, not in five years but in six years, it might be in your cell phone.

[INAUDIBLE]. All right. So let's continue here. Has anybody seen this snake illusion before? Yeah. Does everybody see rotating snakes?

AUDIENCE: No.

RAMESH RASKAR: No?

AUDIENCE: [INAUDIBLE] a head-on [INAUDIBLE].

AUDIENCE: Yeah, it doesn't work [INAUDIBLE].

RAMESH RASKAR: At an angle?

AUDIENCE: Mm-hmm.

AUDIENCE: Yeah.

RAMESH RASKAR: All right. Maybe I can-- how many of you can see it? How many can see and how many cannot see a rotating pattern? So OK.

I've heard this before. Because this is different than those really annoying random dot stereograms, the Magic Eye that a lot of people have trouble seeing. But this is not based on the same-- it's based on a different principle.

It's just looking at the decay in your retina, how long it takes for your retina when it's doing circuits to basically do center surround subtractions. So it's a very different principle. So those of you who cannot perceive motion in this one-- do you also have problems with Magic Eye?

AUDIENCE: Yeah, I can't see anything. I can't see those Eye ones.

RAMESH RASKAR: Yeah, I think Magic Eye is probably-- how many people have problems with Magic Eye? Yeah, I think that's a lot. But so it looks like the two sets don't have any relationship with those type of sets. So how can I take a photo of a scene that has motion printed on a piece of paper and create an illusion of motion? That's the first problem for your project.

AUDIENCE: You could print that out.

RAMESH RASKAR: Yeah, but that's only for this particular scene. I want to take a photo of a car moving, or somebody running. And I just want to take a couple of, maybe five frames of that video and create one photo out of that so that it looks like the person is constantly moving.

I think we can put on the lights again. Let me see this. So all right.

Here was that question, and how can I look around the corner. So this is actually a huge project in our group of, how can you look around the corner? And the way we do it is, we actually use some really cheap devices-- a so-called transit imaging camera, where we use an impulse response of a scene.

We transmit a signal. Maybe I'll use that one. We transmit a very tiny pulse, which reflects off of the door, bounces around inside the scene, reflects back from the door. And it is captured back again at the camera.

And by analyzing these multipass reflections, we can figure out what's inside the room by just looking at the door. So the devices are extremely cheap. We need a femtosecond laser.

That costs about $200,000. And we need a photodetector. That costs about $3,000. But we find some cheaper versions.

And then, we also need a 10-gigahertz scope. That's about $50,000. So after about a quarter million dollars, you can look what's around the corner.

So maybe not in 5 years, but maybe in 10 years, or 20 years, you will have devices which will allow you to look around the corner. And right now, in our lab, you can do it. And this is going to completely change the way you think about photography, because line of sight is almost a fundamental assumption.

We think it's almost-- a lot of us think that's one of the laws of physics. We can only see things that are within the line of sight. But we are not violating any laws of physics. If we're violating the laws, you're welcome to report us. but--

[LAUGHTER]

This is possible. So throughout this class, what you'll see is that the laws that you take for granted and the laws that are shackling the way you think about visual capture and visual displays are just there because somebody taught you that. But you can challenge all those assumptions with modern tools, whether they're sensors, and optics, and modern computational methods.

AUDIENCE: What's the point of the femtosecond laser if it has such a short coherence length? You can't really be using the laser.

RAMESH RASKAR: So the reason why you need an extremely short duration laser is that if you just-- imagine you build a camera. So those cameras, this time of flight camera that Jay was showing-- images of pulse at about 50 megahertz. So it's 29 nanosecond repetition. And light travels how much in one nanosecond?

AUDIENCE: 1 millimeter per microsecond.

RAMESH RASKAR: Yeah, but that's too complicated here. How much in one nanosecond?

AUDIENCE: About a micron.

RAMESH RASKAR: No, much more.

AUDIENCE: An inch.

RAMESH RASKAR: More than that.

AUDIENCE: 100 meters.

AUDIENCE: A foot?

RAMESH RASKAR: A foot.

AUDIENCE: I guess--

RAMESH RASKAR: That's right. That's right.

AUDIENCE: Sorry, yeah.

RAMESH RASKAR: So a bit simple thing to remember. The light travels one foot in one nanosecond. And what about sound at room temperature? How much does it take for sound to travel one foot? You know all the numbers. It's 330 meters per second for--

[INTERPOSING VOICES]

RAMESH RASKAR: --sound and say that's 10 to the 8 meters per second for light. But that's too complicated to think about. We want some simple rules of thumb. So--

AUDIENCE: 10 milliseconds.

RAMESH RASKAR: 10 milliseconds for one foot? It's one millisecond, right? So very easy to remember. Light travels one foot in one nanosecond.

Sound travels one foot in one millisecond. So light travels one nanosecond. Sound travels one millisecond. Unfortunately, there's nothing that travels in one microsecond.

If somebody can, with a new physical propagation, physical propagation channel that travels one foot in one microsecond rather than milli and nano, then you'll see a completely new range of applications. So you have electromagnetics, P-M spectrum. That's one nanosecond per one foot, and sound, one millisecond for one foot.

So sound is too slow, and light is too fast in almost everything we want to do. And so here, we want something that's even faster than traditional light propagation. So in this room, if I just send a beam of light, very narrow pulse, by the time it goes to that wall and comes all the way back to me-- let's say this is about 20 foot-- it's going to take me about 40 nanoseconds to come back.

On the other hand, if I just want to see where somebody is within a couple of feet of that wall, then I need to start measuring in picoseconds-- not 10 to the minus 9 nanoseconds, but 10 to the minus 12 seconds. And that's why you need to do this extremely fast to be able to do any-- so a traditional time of flight camera will just look at this door.

And that's it. There's nothing more you can do. But if you want to look at the reflections, then you need to be able to resolve time at a much higher temporal resolution.

So femtosecond lasers are not something-- they're not exotic. They're used in OCT. They're used in two-photon microscopy. And they're used in a lot of their applications, but still not consumer applications.

They're used in medical imaging. And once they become solid state and easy to carry, we can do that. A lot of LiDAR also happens in not femtosecond, but nanosecond ranges with very high power.

So if you can look around the corner like this, what about looking around a beautiful artifact like this one? If I have a bottle-- where's my bottle? If I, as a human, when I look at this bottle, I look around it. And I create a mental representation of what this looks like.

But if I capture a photo, it's going to be only from a [INAUDIBLE] view. It's trying to mimic my perspective and so on. But my mental representation is actually something like that, right?

So how can we build a camera that takes an object and creates a roll out imagery? Now, this particular object is straightforward. I can just put it on a flatbed scanner.

And I can just roll it. But some other objects are not so easy. This object, for example, which is not completely cylindrical, doesn't have a constant radius around its axis of rotation.

So if I just roll it on a flatbed scanner, I will not get that. So maybe I need a special camera, or I can use my existing camera and use some interesting tricks, computational tricks to create rollout imagery. Another such problem.

And I'm sure a camera company would love to have this feature, as you have all the boring features-- AV, TV, panorama, movie. And then, you have rollout mode. So that would be fun.

And maybe you can do it with an ordinary camera. Or maybe you can do it with a femtosecond laser, if you have a quarter million dollars. All right. So let's do a very fast forward preview of the rest of the class.

And here, I'm mostly going to talk about what's the input, what's the output. I may not go into the detail of exactly how this works, because again, these are teasers of what's coming in the class. And what I would like you to think about during this preview is how this applies to some problem you may be already working on, or what are some parallels with things you already know?

And again, most of these techniques will be about changing the rules of the game. If you have a project where you're tracking your fingers with a cheap webcam and it's not working, because when the light changes or a person with a different skin color walks into the scene, there are solutions here. If you're worried about how to track crowds, there are solutions here.

You want to see what's behind a glass that's murky and diffuse, the solution's here and so on. And then, of course, there are really interesting devices that you could use, and new forms of photography. So here's a really simple example of how we can get started.

So Paul [INAUDIBLE] in '92-- simple idea. Take an object, turn on the flashlight on the left. Turn on a flashlight on the right. And then, you can form this image by combining these two.

How will you do it? This looks like a blue light, and this looks like a red light. Just from these two photos, you want to create this. Yeah?

AUDIENCE: Just mix the channels.

RAMESH RASKAR: Exactly. Just take the blue channel from here and red channel from here. That's it. And it creates this beautiful lighting artifact.

So that's going to be, actually, assignment number 1-- just a warm-up assignment. All you to do is take an object, and take two or three photos by moving your light source. And then, mix and match the color channels to create very beautiful color artifacts.

And this will help you to get your whole pipeline for a sensor going. You will have your own camera. You're welcome to use your regular camera, like even a cell phone camera. But ideally, you should start using a camera that has more manual controls.

And it will get you set up with your MATLAB, or Java, or Flash, whatever you want to use-- C++, Open CV. There's a lot of easy ways to do this. But I would like you to set up your environment so that these type of operations are very easy to do. And that will be your assignment number 1.

Should be very straightforward. Of course, this is what you would do at home. But if you have a couple of million dollars, then this is how you would do it in the Hollywood. So this is a project from Paul [INAUDIBLE] and his group-- really excellent set of work, where they're building different light stages.

And it's just turning on one light at a time, or moving it manually. They have a dome with about 150 lights. And then, they use a high speed camera here, and turn on one light at a time.

So they can photograph this actress under 150 lighting conditions in 150 frames very quickly. And then, I can start from the beginning, and so on. That's one way to do it. The other way to do that is, let's say all you want to do-- insert this actress in a scene that may have been shot somewhere else.

So she's in LA. And you want to insert her on top of a photo that was actually taken in Milan. Now, if you just take the photo and superimpose it, it looks very fake, because the lighting doesn't match up.

So the trick they use is, some guy puts a shiny sphere in Milan and takes an environment map of that courtyard. Then, you feed this image onto this light. So the left corner is reddish.

The light on the left side of this dome is reddish if it's yellowish here, and so on. So for all this 4 pi, you turn on the lights correspondingly. And now, she's bathed in light as if she was in that courtyard in Milan. And now, if you take her photo and cut out and superimpose on that background, it will look more realistic.

And they have done videos where they took these shiny spheres. And Christmas is coming up. So you can pick up your shiny spheres. And you just put the shiny sphere, and move in the courtyard with it.

So you're constantly capturing the environment map as you move along. And in the scene, she just stands in one place. But the dome is lit up by the environment map that was captured from these shining spheres.

So in a video sequence, it appears as though she's walking through this environment. At least, she's lit up as if she's walking through this environment. So these are some of the tricks that are being used currently in major productions.

So *Matrix*-- all the movies that came in the last 10 years or so are using this particular mechanism to create maps that have correct lighting, match lighting. And if you don't want to spend $2 million, back to something really cheap-- to create those silhouettes and so on, you can use a Multiflash camera. So here are some cameras that you can buy.

This is, I believe, a little more. Yeah, this is a little more. So I think it's about $30. And what it does is, let me see if it's-- did the lights go off?

So when you release the shutter, it-- but too cheap to put the film inside of that manually. I'll put that on. I think we're out of battery here, unfortunately.

But anyway, when you release the shutter, it takes four photos by exposing one pinhole at a time. And at the same time, because it cannot recharge the flash that quickly, the simple solution was to actually put four different flashes. So this shutter goes off. Then, this light goes off.

Then, this shutter goes off. This light goes off, and so on. So it takes four pictures.

Now, instead of putting the lights all in one place, if you place them around the camera, you can do something interesting, which would have been, the flash is to the left. You know that we get very annoying slivers of shadow in an ordinary photo.

Now, if you intentionally press the flash-- so you can see it here. So there are all these slivers of shadow. And that is continuous. And you probably see it in your own photographs.

If you place the flash intentionally to the right, then the slivers move to the left. If you put the flash at the top, the slivers are shadows slivers at the bottom, and so on. So by taking these four pictures and analyzing those tiny slivers of shadows, you can figure out where the depth discontinuities are, where the foreground is separated from the background-- not just the whole person and the wall behind them, but also any internal changes.

See, for my hand here, it will create a boundary between my hand and my body no matter how close or how far I am from that. So by doing that, you can estimate all the shape contours. And this is the edge map you would get if you had an ordinary camera.

And this edge map you get with a Multiflash camera. So now, if you have an application-- you want to track a hand or tract a gesture, instead of taking a standard 2D image, if you take a Multiflash image, you get very clean contours. And from those contours, you can build an XC application that will perform very well, even in strange, ambient light.

And it's independent of the color of the foreground object. It only relies on the shadow. So again, if you want to track your hand, you're not dependent on the skin color anymore. So you can play these tricks to overcome the limitations of a traditional 2D camera.

Let me skip ahead a little bit here. Another assignment we'll be looking at is this vertical optical bench. And it's a very nice toy that Andrew Adams at Stanford put together, which is a Flash-based application where you can insert lenses, and occluders, and mirrors, and ray emitters, and so on.

And it can basically do a very quick setup, a very quick optical design of a setup. And what we'll do is, we'll start with this. One option is to start with this his code, his source code, and modify and insert a few more optical elements-- maybe a prism, maybe a grating, and so on, OK?

This will be one option. And as I said, you'll have multiple options for each assignment. So thinking a little bit more about lenses, one concept that we'll come across quite a bit in this class is light field. And now, this particular camera that we saw of the five that Rod was showing us is actually a light field camera.

But that's made up of an array of cameras-- physical cameras. And what we're going to do instead is, we're going to take an ordinary camera and convert that into an array of virtual cameras. So this is an array of physical cameras. But it's expensive.

And so that will take an ordinary camera and convert that into an array of virtual cameras. So this is how it works. In a traditional camera, you-- if the object is in sharp focus, the radius along each of these directions is convergent on a single pixel. So you get a very sharp image of the point.

But any information about the radius along each of these directions is completely lost. So you get 2D image. You have a 3D scene.

You get a 2D image. So it's flattened. The world is flat.

A trick you can do, which was actually invented by Ted Ellison and his student Wong that just left is trying to capture the radiance along each of these directions. So how can you do that? You just displace the sensor a little bit back.

And in front of that, you put a microlens array. This is the same microlens array you use in lenticular displays, those displays that change with viewpoint. So if you put that microlens array, then as you can see, each of these rays is actually incident on a different pixel.

And then, you can capture the variation along each of the categories. Now, why would you care about capturing each of these rays? It turns out that the appearance of the world coming through a lens can be completely described geometrically, completely described by a four-dimensional function, which is this light field.

And that's a very powerful concept, because if you do capture this full representation, then you can do-- that's all you could ever capture. Once you have this 4D representation, you can manipulate that in many interesting ways. So Ted Ellison and [INAUDIBLE] at Stanford, who also has a company now called Refocus Imaging, are building these type of cameras with [INAUDIBLE].

Now, how did the Stanford team do it? They started with a medium format camera with a digital back. And on the digital back, they put this microlens array, where the pitch of the microlens is 125 microns. So pixels are about nine microns. So under each square tile, they have about 14 by 14 pixels.

And so again, going back here, under each microlens, they have a 14 by 14 array of pixels. So what they're going to do now is, they're going to take the 16-megapixel detector, 4,000 or 4,000 pixels, and have 292 by 292 pixel array-- sorry, microlens array. Under each microlens, 14 by 14 pixels.

So at the end, they have this 116-megapixel image, which after reshaping, gives you this 292 by 292 pixel image, 1 under each microlens, OK? So they had given up a lot of resolution from 16 megapixel right down to 292 by 292, OK? But with that, we can do some amazing things.

We can do digital refocus completely in software. OK? So you have given up a lot of resolution. But now, you have complete control over where you can focus.

And as you can imagine, this is the same question you asked earlier, from tests I can also estimate depth. Because depending on when things come in focus, I can assign a depth to each pixel. So suddenly, from an ordinary 2D sensor, I have a camera which has how many virtual cameras? Here, we have five by five. How many virtual cameras here?

AUDIENCE: 200--

AUDIENCE: 14 by--

RAMESH RASKAR: 14 by 14 cameras. What is that, 228? No, 496.

AUDIENCE: But then, each camera is 292 by 292 and--

RAMESH RASKAR: Exactly. It's very low resolution. And so first complaint is, yes, it gives you all this power, but very low resolution. And the argument nowadays against that is, whether you have 6-megapixel camera or a 16-megapixel camera, it doesn't really matter.

We have reached-- we have diminishing returns after six megapixel. So why not use those pixels for capturing some other information? So that's what makes it extremely powerful.

And refocusing is only one. Depth sensing is another. Interaction, dealing with aberrations in the scene-- a lot of interesting things you can do. And again, this is starting with a static 16-megapixel camera. So it's not video rate. But this one is video rate, although it's only 25 virtual cameras.

AUDIENCE: Does the microlens have any way to read what's in and out in a way so that you can combine the images?

RAMESH RASKAR: You could do that. Unfortunately, the precision that you require is micrometer precision. So it's a little bit challenging.

But you're right. I'm sure when cameras were designed in the beginning, they had physical apertures that couldn't be changed. And over time, people figured out how to create variable apertures and so on. So creating these dynamic elements is going to be the key for future cameras.

And people often ask me, what are the things that you're going to see in the camera next? And we already have high dynamic range. Next is color. And we get a lot of journalists asking questions like this-- the future of photography.

A lot of popular magazines with a lot of junk in it, including people who are not credible, the pictures in there. So to me, the answer to that question is light fields. That's going to be the next big thing. If you think about the top five features that will appear in a camera, it's exposure, color-- number 3 is light field.

So we'll see how long it's going to take before we have a full-fledged light field camera as a consumer device. And again, initially, they're going to say, hey, but it's only 292 by 292 pixels. But my guess is that by then, we won't care about the pixels in our business.

So my group was extremely inspired by this work in 2004, 2005. But we thought, this is very challenging to create because you need a microlens array. So we said, instead of using a microlens array, can I just print a transparency at home and create this light field camera? So that's what we did, which is called a mask-based light field camera.

And this is how you do it. You start with a medium format camera-- in this case, a Mamiya. On the digital back, you just remove the IR filter. And it turns out, there is already some glass on top of the sensor, which is about 1.2 millimeters thick.

So you just drop a transparency on top of it, snap back the IR filter. And that's it. For about $2, you can convert a medium format camera into a light field camera.

And the design looks something like this. Traditional camera, traditional sensor, but about 1 millimeter in front of it, you have a printed mask. Very cheap. Using that, we were able to convert this 2D camera into something that captures a 4D function.

And the concept is actually very similar to radio frequency heterodyne. The reason why you can listen to multiple radio stations on a single antenna is because all those stations are transmitting using either amplitude or frequency modulation. And then, in software in your car, you can tune into any one of those channels and decode any one of those radio stations. And what we're doing here is very similar.

We are doing that in the optical domain. And that's what we call it optical heterodyning in space, not in time, where you have the object. It's forming an image on the sensor. But we're going to take this photographic signal, which is four-dimensional, not two-dimensional, and use this carrier, and then create a modulated signal.

So again, for those of you with a communication background, this analogy will work. And then, software-- knowing this carrier, we can demodulate that and recover this four-dimensional light array. So it's possible to do it with a very low cost. Sorry, can you hit the lights again?

AUDIENCE: This one here?

RAMESH RASKAR: Yeah. Maybe the other one. Thanks. And this is a photo that we captured with our mask-based light field camera.

If you zoom in, the in-focus parts are actually OK. Autofocus part has a really strange encoding because of the high frequency mask. But then, software-- and this is how the mask looks like, the printed mask. And in software, it turns out that by applying appropriate signal processing framework, we can decode that and recover it.

So this is a 2D frequency transform of a traditional photo, where most of the energy is in the [INAUDIBLE]. And after applying this very high frequency mask in the optical part, it actually encodes this information-- in this particular case, a 9 by 9 windows of the Fourier transform. And this intentional aliasing or heterodyning allows you to capture the additional two degrees of freedom.

And so the process is very simple. You take this photo, which is about four megapixel. You take its 2D Fourier transform. You reshape that to a 4D function.

And then, you take the inverse Fourier transform to create these 81 virtual cameras-- in this case, 9 by 9. This was 5 by 5. This is 9 by 9.

And the best way to show that is to see how that photo will look like when there is a small parallax between each of the virtual cameras. And again, from these 81 images, you can estimate depth. You can create refocused images and all that-- the same thing that Rod was showing for refocusing from here to infinite-- yeah, I guess infinitely back and forth.

So you can do that with an ordinary camera, with a small change on the [INAUDIBLE]. And all the software is online. And this will not be part of any assignment, but you're welcome to take that up as an assignment or as a project.

Maybe we can get the lights back on and start. What are some other things you can do? As I said, in terms of the desired camera features, it's dynamic range color, and then light field.

In terms of color, Ankit Mohan, who was a scientist from the group-- as part of his thesis, he said, let's think about color as not multispectral imagery, but more like an audio synthesizer. If you have an audio system and you're listening to rock, or jazz, or pop, or country music, you tune your bass and treble accordingly. Or maybe you have a profile so that the frequency profile of your synthesizer is appropriate for that particular type of music.

The same thing should be possible for photography as well. If you are in the woods, you would like to look at most of the green channel to see how the variation of different leaves and all the nature is captured with sufficient variation. Maybe you don't care so much about varieties and looks there. On the other hand, if you are near an ocean, maybe you mostly care about the blue shades and so on.

So what I would like to do-- and photographers do this all the time. They carry a set of filters with them. And if they are looking at broad sunlight, they put one type of filter. If they are on a beach, they put another filter and so on.

What I would like to do is create a knob right on the camera just like an audio synthesizer that says, boost red, suppress green, and create any profile you want. So creating this programmable wavelength would be extremely powerful. So Ankit's project basically achieved that, which it calls Agile Spectrum Imaging, or programmable wavelength imaging. So it's a very powerful concept.

And hopefully, it will appear in cameras as well. You can imagine, this is very useful for medical imaging. When you go to the dentist and they're putting-- what's it called? Enamel?

AUDIENCE: What?

RAMESH RASKAR: To change the color of your tooth?

AUDIENCE: The whitening?

RAMESH RASKAR: Yeah, the whitening. What is it called? Enamel?

AUDIENCE: Bleach.

RAMESH RASKAR: Bleach.

AUDIENCE: Yeah, yeah.

RAMESH RASKAR: And the problem with that is in the dentist's office, everything looks fine. But you go elsewhere, and somebody takes a flash photo and the guy with the fake teeth or bleached teeth-- it looks extremely different. And that's [INAUDIBLE] because the wavelength profile of a flash is very different from the profile of tube lights.

So what doctors would like to do is see the neighboring teeth under all different lighting conditions so that they're still matched, for example. And this is true of-- I did show you the vein viewer, where you want to see the veins. And depending on oxygenated or deoxygenated blood, the hemoglobin you can figure out which veins should be used to poke the needles.

And again, that can be looked into very narrow wavelength. So they don't-- in that case, they might know. But in different applications, they may not know probably which wavelength you should be looking at.

And so they by creating this programmable spectrum camera, you can do, again, very interesting things. So we'll be looking at that. And glare is another challenging problem, right?

If you have bright sunlight, it's going to be glare. Sometimes, it's for artistic effect. Sometimes, it's just annoying.

So can you take a photo that has these concentric rings because of glare and either boost the glare, create some cheesy effects, like it's a rainbow transition here from blue to red, or actually suppress the glare-- again, all from a single photo? So it turns out, glare can also be captured using a light field camera if you have a bright light such as this-- let's see. This is a bright light. And this is some other scene.

The bright light will create a sharp photo. But because of interreflection, the Fresnel reflection in the lenses also create a glare effect and contribute to the wrong part of the image. But again, what we will learn in this class is that by doing this 4D sampling, you have complete control over the lens glare and certain types of glare.

Again, light field concept for a camera array-- Stanford and Professor Mark Levoy are the world leaders in thinking about light fields and light field cameras. So they built this amazing camera array, the electronics for its optics, and so on. In this case, I believe about 51 cameras. And then, they can do very interesting things.

So here is a scene. We have about 51 cameras looking at the scene behind the bushes and trees. And this is how it looks. Focus on that part.

And by doing refocusing, you can see what's behind those trees. So this is just-- you're just doing virtual refocusing in the scene. And by doing that, by using an extremely large aperture, you can see what's behind these bushes. And it's pure refocusing. You can do additional computational techniques to recover what's behind the trees.

So a lot of things in computational camera and photography are really about magic. How can you look around a corner? How you can look behind the trees? How can you look inside the body, and so on? So that's why I like this field.

It's like magic tricks. And once in a while, you come up with your own magic trick. And some other times, people show you a magic trick.

And different people figure out different ways of achieving the same magic. And that's why it's such a vibrant new field. The way synthetic aperture works is the same way-- if I hold a needle next to my eye, not poking my eye but just right next to my eye, then if I focus on the needle, I'll see it.

But if I focus far away, then this needle just becomes blurred. And this does not occlude what's behind. And the same concept is for synthetic aperture. It's used in radar. It's used in astronomy.

And all you have is an array of receivers, whether it's antennas, or whether it's microphones, whatever it is. In this case, cameras-- array of cameras. And if you have a very, very large aperture, then a point that was occluding some point behind ends up being extremely blurred. So it does not impact what's behind that.

If you have a very narrow aperture camera, you cannot do that. But you can do that if you have a very wide aperture camera. So we saw that. OK.

So what about medical imaging, such as computer tomography? And we're jumping from photography to tomography. But both are recording.

One is recording light. The other one is recording slices. It turns out you can use very similar principles.

So what's happening in tomography? You have an X-ray source that's emitting in an omnidirectional fashion. And you have detectors here.

Should I use this screen, because most people are-- which screen is better for the majority of people? That one? Sorry.

AUDIENCE: Yeah, that one.

RAMESH RASKAR: All right. I'll use this one. Sorry, Rod [LAUGHS]

AUDIENCE: It's fine.

RAMESH RASKAR: This one was convenient for me. You have an X-ray source. And you have an array of detectors. And basically, when you put your head inside this X-ray under tomography, CAT scan machine, the X-ray source spins.

And the detector moves in the same direction, right? And let me show you a video of how this actually works. And let's start from the beginning.

All right. It's been opened up. So you know what's going on. That's what's happening. Imagine your head is inside that.

[LAUGHTER]

All right? That's what the CAT scan machine is doing. And while they put an eyepatch and you're just resting inside, it's basically the a engine.

It's totally crazy, OK? It's totally unnecessary. We are the 21st century.

[LAUGHTER]

We are in the 21st century. And we're building these devices that are based on 40-year-old principles. It's unbelievable.

It's totally ridiculous. What we need to do is completely rethink how this imaging is done and use new computational methods to overcome these totally bizarre, multimillion dollar devices. So that will be one of the things we'll be learning in this class-- how we can take principles from signal processing, photography, scientific imaging, and mix and match them to build new things.

Both were tomography-- another very interesting problem. All you do is, you create your drill hole. You put explosives here. And you put sensors here-- or in this case, a [INAUDIBLE].

And then, you fire off these bombs, effectively. And based on how long it takes for sound to travel to these detectors, it tells you what the density of material-- whether it's rock or oil, or other types of formations. And from that, the oil companies can figure out whatever oil there is, and where it is. So you can create a 3D map of what's inside.

Again, multibillion dollar situations. Microscopy, deconvolution-- also used in photography, also used in a machine vision computer, which we'll talk about that. Coded aperture imaging-- an idea that was used in astronomy, because astronomy-- you're looking at gamma rays and X-rays.

And you cannot build lenses for them to form an image. So you can either create an image with a pinhole of the sky, or you can use a coated aperture so you can collect more light. Now, that idea-- we'll learn about that-- can also be used for photography. So what our group did was, we used a coded aperture in the lens.

Instead of having a clear, disk-like aperture, we put a-- you can barely see it. I'm sorry. It's supposed to go this way. [CHUCKLES] You have a coded aperture, which looks like a crossword puzzle-shaped mask.

And from that, you can take a photo, again, which could be out of focus, and then digitally refocus that. So this is autofocus photo. This is in-focus photo.

And you can capture a glint in the eye, or even a strand of the hair. Again, very minor change in the camera. And originally, we were talking about successful biological vision. If you think about the simplest possible biological vision, which is a single pixel detector in a world, right? Just a single pixel.

It's in muddy, marshy waters looking for food, or maintaining its orientation. It doesn't need a full-fledged camera, it just has a single pixel detector. It just knows if there's light, there's no light, if there's light, how much there is this. So you know when it's dark, when it's day, when it's night, which way is more light, which is less light.

But even that single pixel detector has some very interesting optics in front. It has this very intriguing shielding pigment in front of it. Can anybody guess what's the reason for that? How does it benefit to have some random pigment that's blocking the light from different directions?

AUDIENCE: For orientation?

RAMESH RASKAR: For orientation. So if this worm wants to maintain an-- if there's a light source and if the sensor was hemispherical, then if this worm moves a little bit, there will not be much change. But if they have a very high frequency pigment, if this worm moves even a little bit, there'd be a big change in the lighting, right?

It's as if you have a very high frequency mask and you're looking at the sun. As you move, the light goes up and down. So the worm knows that as long as it's maintaining the same level of light, it's maintaining its orientation. That's all it needs to do.

AUDIENCE: Does it also increase the effective dynamic range of its sensor?

RAMESH RASKAR: It's possible.

AUDIENCE: As a neutral density filters along certain directions, and--

RAMESH RASKAR: It's possible. It's possible. Maybe when it's looking at one direction, it's too bright. So it tries to block that.

We don't know. And if you read this book, there's a beautiful book called*Animal Eyes* by Land and Nilsson. We have a whole class on animal eyes. I think it's the seventh class, if I remember correctly.

And we'll discuss all different types of animal eyes and why they do it, whether it's eagles or land creatures, underwater creatures, worms, and all that. And frankly, most of these biological visual systems are based on hypothesis. And they're verified in a very, very scientific, but at the same time, very primitive equipment.

And one of the great projects would be to take some of these worms and put it in a controlled lighting setup to really verify if this is how they work. That'd be a lot of fun. I'll provide the worms. Don't worry.

[LAUGHTER]

So the squared aperture is somewhat similar. Let me skip over this part, because we talked a little bit about how these mask. All right, wavefront coding.

This is a concept that was invented by Cathey and Dowski in 1995 for shaping light or shaping the wavefront of incoming light. So you have a traditional imaging setup. You have an object.

You have a sensor. You have some lenses. What they proposed is placing an additional optical layer in-between, which is not a lens or a prism, but has variable thickness or variable refractive index, OK? The simplest way to think about that is, the light in the top part might travel at one speed.

Remember, if you have a glass with different thickness or different refractive index, the light is going to slow down. And when it comes down on the other side in air, it can go back to one foot per?

AUDIENCE: Nanosecond.

RAMESH RASKAR: Nanosecond, very good. But before that, it's going to travel slightly less than one foot per nanosecond, right? If it's 1.5, it's going to travel 1 by 1.5, 1/2 foot per nanosecond.

So anyway, so by adding glass of different thickness or different refractive index, each of these rays are going to be slightly out of phase with each other. So when they combine on the detector, they will interfere either constructively or destructively. And from that, you'll form new images, right?

So this is how they explained it. And if you read the papers, unfortunately, they are very difficult to understand. And what you'll realize is that instead of going into the math for your optics and so on, in this class, we will use very simple ray diagrams and understand how this works in a very visual manner. OK.

So basically, what wavefront coding camera does is, in a traditional camera, rays converge to a single point. And you get a sharply focused image. So if the sensor goes in and out of focus, you get a large blur. In case of wavefront coded camera, actually, you never get a sharp spot.

What you get is basically, think of taking a lot of straws and all of them converging to a single point, taking them and then twisting them so that they go out again as the straws. But in-between, there's a part where all of them are-- the cross-section of them is roughly-- OK? And by doing that, it turns out, for a sufficiently large depth range, the defocus is equivalent. And we'll study this in detail, and how this works, and how you can use the same techniques for new types of photography and scientific imaging.

Now, this is also a very hot topic in night vision goggles, by the way, where they want to wear night vision goggles that when you look far away, it's very clear. And again, night vision goggles have very large apertures. And when you look closer, if you want to read a map, for example, it should still be in focus. So how do you create a passive device that can focus on infinity and very close up at the same time?

And they have been using wavefront coding as well. Wow, I can barely see this. All right. This is a project called "Decoding Depth via Defocus Blur."

This is from Colorado from Rafael [INAUDIBLE] and [INAUDIBLE]. And this is very counterintuitive. If you take a point light-- and I wish I could do this experiment. I can do it here, but it'll take some time. If you take a point light and form an image on this particular guy--

AUDIENCE: Do you want a little flashlight?

RAMESH RASKAR: Yeah. Thanks. So as you can see here, that's the image plane, right? As I move it in and out, its shape is going to change-- the spotlight.

Now, imagine-- and that's what happens in the bottom part of the image. As you are in sharp focus, you get a small spot. As you go out of focus, you've got a bigger disk.

But the kind of optics this group designed, when you're in focus, you see two spots that are left and right of each other. And when you go out of focus, these two spots rotate one way or the other way. If you are closer than the other focus, it rotates in one way, if away from other focus, it rotates the other one.

So this rotational point spread function is a very powerful concept. And right now, they're using it in microscopy for resolving fluorescent beads at nanometer precision. But it could also be used in photography.

Nobody knows how to do it. This could be one of your research projects. And we can get some of these prototypes from our colleagues. So it's a lot of fun to play with.

Another interesting thing we will be looking at is the relationship between Fourier optics and ray optics. Now, in high school, we were taught, there is particle and wave duality, and a lot of confusion. And in high school, we have to answer questions of what phenomena can be shown in particle way, and what can be shown in using the wave propagation model.

And one of the standard answers was, oh, if there's interference or diffraction, then it can only be explained with Fourier optics, not with particle nature of light. That's too simplistic. It turns out that you can really show that there are the duals of each other and explain diffraction, and interference, and all these mechanisms using purely ray propagation.

And so with Sadek-- his name is not here-- Sadek O and Josh [INAUDIBLE], this is-- for mechanical engineering, we have a project where we have created a so-called augmented life field, which can actually support all the wave optics effects as well. So traditional light field that we just described earlier, where you can capture with an array of cameras and so on, can do this position angle representation and the four-dimensional representation.

And at that time, I made a claim that if you can capture this 4D incoming function, you have captured everything-- everything-- geometrically that has come through the lens. So it's a complete representation of the light. And some people would say, wow.

But then, you're not capturing the phase. And you're not capturing all these other things. But it turns out that using the same exact setup, you are also able to capture phase. The Fourier representation actually includes the amplitude and phase of the incoming wavefront.

And so simply by using different mathematical terminology, we have augmented this light Fourier presentation to also represent, to also model wavefront effects. So we'll study that a little later. What about photographs like this? If you look at this heater, it's creating these beautiful streams of hot air.

And also, this lamp from the lamp ship. Again, we are trying to visualize that cannot be seen with the naked eye. So this is known as the Schlieren photography, which is, again, looking at very minor changes in the optical path.

In this case, because of heat, on a hot day on a highway, you see the mirage. But that happens at extremely high temperatures. Now, even at not-so-high temperatures, you can actually capture this mirage and create a very beautiful focus.

So this is student photography. And we'll study that. Polarization is beautiful. You may have used polarization for taking photos of the sky, or on water.

But underwater photography can be dramatically improved with polarization. So we'll study that. We'll also study some new type of sensors, not just two-dimensional sensors, but these other sensors-- a single-pixel camera, compressed sensing.

How many of you have hardware compressed sensing? Right? There's a lot of hype about it. And again, if you read the papers, sometimes, it's very difficult to follow what exactly they're trying to say and how it's done.

But as you will see in this class, you'll get the whole idea of compressed sensing in less than five minutes. And we'll go through what works, what doesn't work. And actually, Rohit is working on a project which hopefully will show that compressed sensing is actually not such a great idea for imaging. But it is good for something else.

And we'll study what that "something else" is when he and Rashad's project is reaching some mature stage. So very exciting ideas there. So the original idea was to create a single-pixel camera.

In creating a megapixel camera, you have a single pixel that's going to take a coded combination of incoming light. This may be a scene. It's being focused on the whole reflective mirror array.

And you're usually going to flip this mirror on and off. So the light you collect will be a product of the scene with this vector. And if you change these flips, you get some other sum. You're going to make a linear sum of incoming light and predict a million such measurements. You can reconstruct a megapixel photo.

But the claim of this group from Rice is that you don't have to actually take a million readings. You might be able to get away with only 10,000 or 100,000 readings. So you can capture a megapixel photo with possibly only 10,000 pixels of a camera.

And we'll look into that and what part of that statement is true, and what part of the statement requires more analysis. And you'll also see that this can be used in a lot of other situations. So for example, my group has built a strobing camera that can be used for laryngoscopy, where in a laryngoscope, you use very high speed strobing to slow down to visualize the motion of vocal force.

But we just came up with a new method that is dramatically simple and uses dramatically less light. So you don't burn your throat when the doctor is looking at your vocal force. And that's based on compressed sensing.

So let me stop here, because we're almost at 4:30. And we'll come back. We'll come back and look at some of these other projects.

So these are the kind of project assignments you will see-- relighting, the first one that's already described. Dual photography, where you can read your opponent's card. Virtual optical bench, light field capture, high-speed imaging, thermal imaging, multispectral imaging, range imaging, and so on. And then, a completely open-ended final project, which you can choose in any area.

This is the first assignment. The instructions will go out on the Stellar web page. Please make sure you have the sign-up sheet. Has it gone around?

OK, please make sure your name and email is on the sign-up sheet. If you don't get an email from me by Monday morning, please send me an email, which means that I could not read your email address correctly. So yeah, these will be some of the assignments you'll be doing.

And this one is due on September 25. And every class starting from next class, we'll have a volunteer who takes notes for the class and posts it, because a lot of our discussion is going to be on the board and so on. I'll send out specific instructions. So we need a volunteer for next week. You want to do that?

AUDIENCE: Yeah.

RAMESH RASKAR: What's your name?

AUDIENCE: Sam.

RAMESH RASKAR: Sam. All right, Sam is going to volunteer for next week. And we'll decide who's going to do that another week.

[SIDE CONVERSATIONS]

AUDIENCE: [INAUDIBLE]?

RAMESH RASKAR: Sorry?

AUDIENCE: Do you have a [INAUDIBLE] from the [INAUDIBLE]?

RAMESH RASKAR: Probably not, because most of the people are going to be-- many people, not most-- but many people are going to be listening. So as far as I know, it's not going to cross the limit. But if there is, there's a cap.

I don't think it's going to reach the campus. There's a cap of authorities. I don't [INAUDIBLE].

[SIDE CONVERSATIONS]

And we'll see some of these demos next time.