PROFESSOR: So looking at some things here, right, x and theta and how they can be represented. But you can create some really beautiful images out of that. You can even create-- some of my slides are not on here. You can create some really interesting pictures like this.

So if I-- this picture is made in that x-theta space. The coordinates are a little bit changed here. This is the slide from [INAUDIBLE]. And it's a robot that Andrew Adams built at Stanford just using LEGOs, and you can do this for your assignment as well. You can build a LEGO-based Gantry that can shift and take pictures. For your assignment, you're just doing 1D translation, but you could even do 2D translation.

And then imagine if I put a camera here, and I simply translated it. And now we're in flatline. When the camera's here, I will take a picture and put it over here, this much. There I'll put the camera, and I'll stack those images.

So for example, when I'm-- so if you look at this particular green object over here, you'll see that it creates, actually, a line because when I start translating it, it gets that line. On the other hand, this particular white sphere ends up mapping, which is really strange.

So this is the visualization of the same space that you saw earlier, which looks maybe simple and very boring. OK. So once you start putting interesting objects, that's how [? your light ?] [INAUDIBLE].

And this is [INAUDIBLE]. Here is-- t and theta are the same as x and theta. And the appearance of this room through this window, all you've got to do is look at this wall through this window. You cannot go inside. So let's see from outside. Everything that can be captured about the scene through this window is represented in this 2D light field.

So now if I want to create any refocusing effects, any depth of field effects, all the information is available here. And if I just reduce the pixels from here and take projections and so on, I can create effects. This-- we're going to twist your, twist your mind around it, and then it starts making sense. For those of you who have a computer vision background, this is also known as AP polar plane images, or API. OK.

So let's go back to some of this concept we were talking about earlier, where you're going to represent each part of the lens as a different camera. So [INAUDIBLE] subsections [INAUDIBLE], except we know that if all the focus at infinity, I can just take the sum of all these images, and I'm done for average of those images.

But if I were to focus closer, then I cannot just take the average. I will shift and add it. And what this prism is doing in the real world is doing the shifting for you. So let's go back to this example of the discussion of shifting and adding.

So I'm going to compare this with the five cameras' picture. [INAUDIBLE] point here from the first camera. The main focus, here. The second one, it was there. Third one, it was there. Fourth one goes here, fifth one goes here. And that's my pinhole camera.

So as you can see, if the point is close by, the coordinates here are not the same. Here is near zero. And here's further up from the center, and here's further down, down from the center. That's all.

So when I have to bring these five images, I need to shift these images so that [INAUDIBLE] overlap on top of each other so that this point will be sharp focus. The lens does that for you automatically because the prism, if I put a painting here, and there is just a pinhole here-- this is ABC. This will come back as ABC.

OK, but this one here would have created an image that looks like this. But because of the prism, it rotates the image. And then this thing shifts down. So they're on top of each other. And because of this part of prism, it shifts. It will have-- the image will have been created here, with ABC.

We'll shift just a little bit and so on. And for the other one, the image would have been formed here. We shift up. So the prism does the job for you, basically, of shifting these images. So it's a very simple way of thinking about how we can emulate a very large lens by [INAUDIBLE] [? cameras. ?]

So the top cameras here should shift down in the images, and the bottom cameras should shift up. And by doing that, you can focus on your wave [? marker, ?] except that, in case of a traditional lens, once the photo's taken, you're done. In case of an array of cameras, you can take those pictures and shift as much as you want and add them in interesting ways. And that's why light fields are so powerful.

So instead of using now an array of cameras, we're going to build landslide arrays or some other fancy optics so that we can directly capture this type of images in a single slideshow. Now, this is three ways of creating-- yes.

AUDIENCE: So would these [INAUDIBLE] be able to [INAUDIBLE]?

PROFESSOR: This is conceptual.

AUDIENCE: [INAUDIBLE]

PROFESSOR: Yeah, you're right. So when you think about the Stanford camera, you want each of these cameras to have [INAUDIBLE] appropriate. So everything is in sharp focus, which means it's a pinhole camera.

So conceptually, yes, it's a pinhole camera. But in the real world, you cannot have a pinhole camera. So--

AUDIENCE: You still try to keep [INAUDIBLE].

PROFESSOR: They got to keep this aperture small. But-- and if you use a really cheap camera, like a cell phone camera, they usually have a pretty large depth of field anyway. So they're as close to pinhole as you can get.

OK, so basically three ways of capturing, light through it-- something that exploits shadows and can hold a ray, 1908-- something that uses [INAUDIBLE] [? array. ?] Also, early 1900s and so-called [? heterodyne ?] of 100 years in my group, OK?

So 200 years to come up with a third solution. Think about how you can-- a stop aperture is basically a pinhole with a prism. So I can take each of these pinholes and just-- so you have this, you have this.

And the third one would be-- I can just put a pinhole. And in front of this, I can put a prism. In front of this, I'll put a shallow prism. In front of this, I'll just put a piece of glass. Here I'll put a-- all right?

So [? lens ?] is basically out of [? pinholes ?] with the setup process. And as I saw, if you have two corners of this pinholes, they [? measure ?] before the [? good ?] [? luck ?] sensor. But unless there's a [INAUDIBLE] of just bending light [INAUDIBLE]. OK?

So this figure is from Lippmann, I believe in 1930. Sorry, this one is Lippman. This one is Ives in 1933. And he said, OK, if you really want to capture each of the [INAUDIBLE] individually, you can just put a pin for that.

So [INAUDIBLE]. OK, so if I put a pinhole out of there and take a picture-- so now, let's [INAUDIBLE] just a pinhole error. Then you end up getting a picture that looks like this. And if you zoom in on that, you'll see that under each pinhole, it will create a disk. So remember, we have a lens slit. And we had a lens slit here.

Instead of that, we just want to have a [INAUDIBLE]. And then [INAUDIBLE] here's a pinhole error. So that's an example for just a pinhole error. So this is all solution number one. Here's our solution number two. And solution number three we'll see in a second.

And it turns out, if you think about a glare in a scene, like a lens flare, a 2D image might look like that on the top right. But if you zoom in, it turns out, in the 4D space, the flare manifests itself as the bright spots in a [INAUDIBLE] image. So you can just go through this image and eliminate all the bright spots, and you can get rid of lens flare.

You know, so you can remove the outliers. Just do some kind of immediate filtering in our neighborhood, and the lens flare [INAUDIBLE]. All right, so the second way of doing it is by using lenslet array, which we already saw.

So I don't want to repeat that again. And now comes the question of, what happens to points that are in focus versus out of focus? And this is a very key concept. You have to understand.

So let's say I have a lenslet array. And I have a point that's in sharp focus at one of these lenses. And same situation here. The point is in sharp focus at this pinhole.

Mike, we'll go back, and I'll get this here. But listen, this is a red crayon. All the rays are red. And when they come back to this pinhole, this whole thing was red.

So that's why, in this part of the image, where in the original photo, with the original camera, the red crayons are in sharp focus. The whole disk is there, which means that the whole-- every part of the lens is getting a red ray from this point of view.

But let's go to a part that's actually out of focus. So I believe this one is over here. So that's on the boundary between maybe a yellow crayon and a blue crayon. So even in the original photo, it's out of focus. It didn't have [INAUDIBLE] focus. And here you can see part of each blob under a lens flare is yellow, and the other part is blue.

OK. Hold that up. So we have a situation where we have a-- it's out of focus. We have a yellow and blue, yellow and blue. But that actually autofocuses.

So if I start shooting the rays on that, [INAUDIBLE]. If I just take the one from the yellow-blue boundary, for this part, the light is in focus closer. OK, it's in focus closer. And the image that shows.

So for the bottom part here, all these rays are going to be [INAUDIBLE]. But this part here is not [INAUDIBLE]. And for the blue part, again, we shoot the rays. And what we realize is that-- because this thing [INAUDIBLE].

If I had a yellow cone coming through and a slightly offset blue cone coming through. And because they were not in sharp focus, they will contribute to one people so that part of it is yellow, part of the block is yellow. Part of the block is-- now, let's go in the-- this one is the [INAUDIBLE].

Let's do the one that's farther away. It's too bad the image [INAUDIBLE]. All right, let's look at this one. [INAUDIBLE] example. So if we just look at the green and the one that's kind of blackish next to it.

So this one was very easy. The blue was to the left of the yellow in the original picture and even inside each blob, the blue is to the left of here. But if you see this one, the green's to the left of the dark region. So it has [INAUDIBLE].

So here the left recording is maintained. Here the left recording is switched. Can anybody tell me what's going on here? Just [INAUDIBLE] exactly because in front of the focus plane, the left [INAUDIBLE] remains in the same [INAUDIBLE].

When you are behind the focus plane, the order is switched. And that's why you cannot simply take this image and reconstruct the original high frequency, high-resolution image. There was a lot of processing for this to be able to recover that.

So it's a lot of fun to look at these images. There's all this information that's encoded, the four-dimensional information. OK, so here's our solution. Instead of placing the lenslet array, we're going to place a mask [INAUDIBLE] array, but especially a printer mask.

You can just take a medium format camera, such as [INAUDIBLE]. Just remove the [? IR ?] protective glass. And then simply plop a film on top of the medium-format camera sensor. Put back the put back the protective glass, and you're ready to go, and you have a [INAUDIBLE] camera.

So in this case, we are not putting a pinhole error, but so I'm going to go through the slides [INAUDIBLE]. Instead of putting a pinhole array, we're going to place a different type of mask.

We captured that. And a few concepts that we had to look at before we get there is that there's this concept of [? conjugate waves ?] that hopefully is in sharp focus, our sensor, then this plane, and the corresponding sensor plane are conjugated with each other.

If I put an object closer, then I must move the sensor further back. And those two are conjugated to each other. And that is defined by the lens equation.

1 over f equals 1 over u, just 1 over u, OK? Now, the key concept here is that a lens copies a light film from the outside world to the inside world. What does that mean? If I have-- so basically, what we were doing, we were assigning some coordinate system. We called it a taa. And we call this one the x.

These are two-plane [INAUDIBLE]. Now, there's a little bit of fudging going around here because theta here doesn't really correspond to absolute angle, but a coordinate on this axis.

So theta is 0 here, and it says plus 5 here and minus 4 here. And x [INAUDIBLE]. And this is over x, the relationship between any point here and any point here if I [INAUDIBLE].

Now, it turns out-- let's say I have two points here, a and b, and two points here, a and b [INAUDIBLE] here. I can also assign a new plane here [INAUDIBLE]. This light stream, which means this x-theta relationship is maintained in this theta-x relationship.

So if I take array here from-- I'm going to assign this also-- let's say we're operating here. And I'll assign the coordinate of 1 over 1,000 and get us a coordinate of 1 over 1,000.

Let's look at some real numbers. Then if I shoot a ray from either 200 plus 4, I can guarantee if the plane wasn't conjugated, that the form will map back to [INAUDIBLE]. It'll map to the [INAUDIBLE].

And of course, you can have 200 mapping to 3 here. 3 will automatically map it also. And that's only because this painting is in sharp focus [INAUDIBLE]. This notion of conjugation means that the light stream of this x-theta array is mapped to this theta x, for this.

So I can take [INAUDIBLE] I can shoot all the rays here. If there's a point here, then I shoot rays out of them. The same point will be sharp focus here, and all the rays will [INAUDIBLE].

OK, if I have two rays that are starting in the same theta direction, they'll come out of here in the same theta direction, OK, and so on. So basically, I'm creating an exact replica of light through here [INAUDIBLE].

And this is possible only when the lens is extremely good [INAUDIBLE]. So it might do a pretty good job, like, it wants to do that. So pretty good, do a good job of popping it in the center, but not at the edges. Or it might be good for one [INAUDIBLE], but not for others and so on.

And this notion of popping the light switch makes the lens very unique because we know that the appearance of the world is 4D. It completely describes the appearance of the world. And what the lens has done is optically, it has copied the appearance of the world faithfully to something that's inside.

And now we just need a good-quality sensor to basically take the hologram of what's out there. And I'm using the word hologram to indicate the 4D appearance of the window into the world. So this window that we have here is exactly copied 4D.

So anything that's-- behind this window, it's faithfully reproduced over here. Unfortunately, in a traditional camera, you take that window and map it back to a 2D image. So the problem is the sense that the optics is doing its job of going from 4D A. So 4D A, right? 4D A.

But the sensor does a terrible job of maintaining that 4D [INAUDIBLE]. But if you have these three solutions, you can recover this 4D image. So let me just explain this concept one more time of how the rays, where there is sharp focus versus auto focus, look differently.

And I promise you that after we understand this concept of rays and ray space and all that, everything else-- out of focus and depth of field-- will become extremely easy to understand. So just see how this works out. So now let's see.

We have a point here. OK, we can either represent the light through here, or we can represent it here. Doesn't really matter. Now, here what we have is a point that's emitting light in all directions.

How does the light actually look like? There's a point x, has [? emanating ?] light in all the directions. Now we know that the lens is going to make an exact replica of that inside.

So when I come here, I have a particular x, and light's coming from all different directions. So what's coming at the plane is exactly that, OK? From this, how do I form an image?

How do I realize that, if I integrate the radiance along each of these rays, I will get the intensity [INAUDIBLE]? So we're going from a 2D world now to a 1D world. In the general case, we go from a 4D world to a 2D sensor-- 4D representation, 2D sensor. Here, we have 2D light fill and 1D sensor.

How would you do it? So with that, I'm just going to sum up everything that comes out of here now, as I'm learning all these values and sum them up. And mathematically, that's just taking a projection.

I'm just going to take this 2D world and flatten it to a 1D world, or you could call it light integral. So imagine I have all these tennis balls here, and I just drop all of them under the force of the aggregate. The number of balls that would come here would be the intensity of the prism and [? is black ?] everywhere else.

Now, what happens when something is out of focus? If it's out of focus, then the lens is doing a good job of transferring the 4D here, 4D here. But it did the right job for this plane, not for this.

So how do we represent the new space of light field? This [INAUDIBLE] simple. There's just one [INAUDIBLE]. They have a bunch of rays and want to represent that.

Previously, the light was reaching only one x point. Now it's reaching an array of x points. And for each of the x points, there's only one direction in which [INAUDIBLE]. So there's still a light once the x is missing.

OK, so I have a bunch of different points here. And for each of them, there's a directional on which the lines are. So there's the notion of shear. We have a straight line with shear development.

And now, how do we compute the light? How do we compute the intensity from the surface? Again, the same operation of, imagine we had all this tennis balls here, and we're going to let them drop?

If we let them drop, we're going to get intensity that goes over a set of pixels rather than just one pixel. And that's the captured photo, you'll see. When something is out of focus, you don't see a sharp point, but you see a blurred set of lights.

So again, we start from here, exact copy. For a given x, we have all the thetas, we project, we get a sharp point. That's the intensity of the captured photo. If it's out of focus, the set of rays can be represented on a slanted line now.

When we project that, we get the integral. And that will be [INAUDIBLE]. Is this [? here? ?] So now when you're thinking about, not only thinking about doing assignment, and you have a set of photos, and you want to do a refocusing, you can think of it in multiple ways.

If I stack on the photos-- OK, so remember, taking a picture with a of cameras is the same thing. So I'm going to put an a of cameras here. And the corner of the camera is theta. And the framework of the camera is x.

So I can take a picture of the first camera, and I'll put that as the top row of this. I'll take a picture with the theta equals 2. I put it over here. Sorry, theta equals 5. Over here-- 4, 3, 2, 1, 0. If I'm nine cameras, I'll put them over here.

And if I just want to take a picture that's focused at infinity, I'm going to start all those pictures and just take the sum while I'm holding the direction. And that will focus at infinity.

If I want to focus closer, then I should just sum them up as they are. But I need to slightly shear them. So I'm going to keep the center camera fixed-- center camera image as this. I'll take the theta equals 1 camera and shift it to the left by 1 pixel. At [INAUDIBLE] equals 2, I'll shift it by 2 pixels, 3 pixels, 5 pixels, minus-- sorry, minus 1, minus 2, minus 3.

And the one over here, I'll shift that plus 1, plus 2, plus 3. If I sum them up, I'm actually focusing on a different plane. So this is the main concept behind refocusing, that it's shifting and adding.

And you can think of that within a single lens, or you can think of that using [INAUDIBLE] cameras. Is it clear so far? So next time you're taking a picture, think about how much work the lens is doing.

It's copying the light field, right? It's capturing the hologram of what's out there. This is using the popular terminology. It's capturing the hologram of what's out there. It's recreating that hologram very close to the sensor, and all that the sensor is doing is just recording that as a 2D image because the sensor doesn't have an ability to record in a traditional sense a 4D hologram. It can almost sense a 2D image. And explain all of that at the speed of light.

Isn't it great? And that's why computational photography is exciting because there's some work you can let the computer do, and there's a lot of work you can just let the physics do for you. But when the physics does it for you, it happens at the speed of light with almost no additional cost. So doing this core design of the physical device and the computational device brings in the real power in computational cameras, computational photography.

All right, so let me switch to some other things I wanted to show you. By the way, is this all clear so far, this part here? This will be on your exam, remember. This is a fun part.

All right, so now that you understood all these concepts of rays and 4D space. Let's see how it impacts-- we already saw all the focus. But let's see how it impacts some other elements of it.

So let's say you have a point right. You take the photo in sharp focus. It looks like a point. If you take the photo, autofocus, it appears as a disk.

And if you look-- if you saw this resolution chart, then it was blurred as well. And here's an example, to answer [INAUDIBLE] question, where we started with a 2D image and now we have another 2D image, but the blur, the image was done more effectively by this disk. So we took every point, base your disk around it, summed up all those values, and assigned it over there, and which was achieved using the same effect as this.

Right. Every pixel in the world, like this one here, actually is contributing to a disk. The other way to think about this, if I just take one pixel here and go up vertically, it's coming from different points in the world. And that's how we can specify [INAUDIBLE].

So now we have that. I'll come back to the [INAUDIBLE]. But every once in a while, we can create some really interesting-- and this is called bokeh in Japanese vocab. What's the right way to say it? [INAUDIBLE].

AUDIENCE: In Japan?

PROFESSOR: Yeah.

AUDIENCE: Bokeh.

PROFESSOR: Bokeh.

AUDIENCE: [INAUDIBLE]

PROFESSOR: All right, good. And, you know, the exploitation of out-of-focus blurring. But sometimes you can do some interesting things. Instead of keeping the aperture completely open, you can insert a special pattern.

OK, so we're going to place this crossword puzzle shaped mask in the aperture. And now if I take the photo out of focus, instead of getting a disk, the LED that you saw earlier, out of focus would [INAUDIBLE] something like this.

The same 7 by 7 pattern that we have here ends up actually going. [INAUDIBLE] and this whole pattern [? fits in. ?] So you can do some really interesting things with it. So again, photographers want to take pictures that have really beautiful bokeh.

So when it's really tiny aperture, things are to focus over a large depth of field. But with a very large aperture, the background is completely out of focus. So now you can start playing some really interesting [INAUDIBLE]. You can take a scene, which has the tiniest bright spots on a [INAUDIBLE] fashion.

And in your picture, you can, instead of putting the 7 by 7 crossword puzzle shaped mask, you can start pulling some alphanumeric [INAUDIBLE]. OK, so I'm going to take seven different pictures-- one with this aperture, one with this aperture, one with this aperture. And if you just take an additional, those three photos-- so building a picture with this aperture, you get a vertical line for every bright spot out of focus.

For this one, you get, so a vertical line, vertical line, vertical line at a different position. You just sum them up. Every out-of-focus spot will have a little [INAUDIBLE]. So you can-- I believe there's an animation here. So you can see letter eight appearing at all different places.

So let's try to understand what's going on here. So that's with a disk aperture, like a traditional aperture. This is the one with some special aperture. And if you want, you can even say, happy birthday, Jennifer, and it would show up there.

So if you take a picture of all the candles, every candle will say happy birthday. OK, so how is this taking place? It's [? difficult ?] to explain purely as out-of-focus blur. But if you think about what's going on over here, it's [INAUDIBLE].

So we have our lens and, you know, sensor. You have one candle or one bright spot. If you're out of focus, it will create a disk here, right? But imagine you started putting some [INAUDIBLE].

For simplicity, I'm just going to make 1010 [INAUDIBLE]. So I'm going to put a full year. That's open, closed, open, closed. So in this part of the lens, I can go through [INAUDIBLE].

This part of the lens I'll block [INAUDIBLE]. Over here, I go through. So I get over here, and this part of the lens is blocked. The part [INAUDIBLE] here will be the whole system.

That's a little blurry. And if I put a different pattern here, we'll get [INAUDIBLE]. And that's how you can [? code ?] [INAUDIBLE].

How do you think about this in the next data sets? But [INAUDIBLE] is not as simple. There is not just a point, but a whole bunch of things going on here. And all of them are [INAUDIBLE].

If you go to that single space, what we have is we have the x space here, [INAUDIBLE] here. And blocking this part and this part, what I'm saying is that the darkest point have blocked this part here, this [INAUDIBLE].

And this part is open. That's fine. And again, the bottom part is [INAUDIBLE]. So what the optical system is doing here is taking the light from [INAUDIBLE], and it's just deleting all the rays here. It's just blocking all the rays. And then [INAUDIBLE] picture can create a [INAUDIBLE]. That's the [INAUDIBLE].

And then, again, you can play this cheesy trick of taking multiple images with different blur. And because of linearity, which we discussed last time, of flight, you can just take the addition of two images to create an illusion that the aperture actually had those three apertures.

So instead of putting a seven in the aperture and taking one picture, you can take three pictures with each of those different apertures and just take some of that. And that's the beauty of light and interaction of light at normal intensities. You can first take a picture and then take an addition as opposed to adding it in the physical domain and then taking a picture. This is good?

So it's a fun project if somebody wants to try it out, you know, create nice bokeh, beautiful patterns. You can also put an LCD in the aperture so it's not just a film, and it can change the LCD so that depending on the event, you can change the different pattern, and you'll get very interesting effects for certain photos.

OK, so what's going on with the animal eyes? And we'll have a full lecture a little bit later on animal eyes. And compound eyes of animals are also very interesting. These are basically array of lenses.

But there's hundreds and hundreds of [INAUDIBLE]. And this is kind of an artistic rendering of what the creature must be seen, just an area of very tiny images, which is not true as we'll see later. But that's how the rendition looks like.

So there are projects such as [? Tombo, ?] which tries to mimic this concept to basically reduce the thickness of your camera. So if you have a 35-millimeter lens, it will be about 35 millimeters deep. But imagine if you want to create a camera that's 3D [INAUDIBLE].

One way of thinking about that is I can just split my lens into a set of tiny lenses. So we should go back to this diagram over here. I'll have a sensor, and lens flare, and main lens. I'll just get rid of the screens. All I [? have. ?]

And the question is, can you do something useful with that? Any guesses what you can do, what you cannot do? That's the situation. So there is no main lens out here. Every point in the world maps to every block. Just got rid of [INAUDIBLE]. So if you have 50 such lenses, every point will [? be seen ?] 50 times. But each of the image is actually pretty low resolution.

AUDIENCE: Isn't this basically the same as the light-field big rig?

PROFESSOR: Repeat that.

AUDIENCE: Isn't this basically the same setup as the light-field big rig but smaller?

PROFESSOR: Big rig?

AUDIENCE: Yeah. The camera rig.

PROFESSOR: Yes, yes. OK, so that's a good point, right? So let's say I tell you that you have a 16-megapixel sensor. Now, 16 megapixel, in flat [? land, ?] you have 4,000 pixels.

And you can use the photonic pixels any way you want. And you have some resolution, in [? theta ?] [INAUDIBLE] resolution, and x. And if I don't use the light-field camera, how focal decreases.

If I place a light field, and let's say under each lens, I get 10 pixels here. I'm subdividing my lens into [? tensors. ?] And my actual resolution's only 400. So I have lens flares going from 1 to 400. And after each lens flare, our pixel-- that's how it enters this.

So in that sense, my resolution in x is 400 and resolution in theta is 10. And you can see that in the x-theta space. It's this way. So this is only 10. And this is 400.

And it's like this photo [INAUDIBLE] because how we form those measurements. [INAUDIBLE] in a difficult sensor, all of them will go exactly for one. There is no [INAUDIBLE]. So I'll have 400 pixels with no resolution in theta.

In a light-field camera, I could have 400 pixels and theta resolution of [? f. ?] Or I could go in other directions, where maybe this is even thicker and more [INAUDIBLE]. I could do, for example, 200 [INAUDIBLE]. I'll keep going. So I'll start it with something good, 10, 5, [INAUDIBLE].

If I could [INAUDIBLE] fade lens-- I'll give you a hint-- what you end up getting is-- [INAUDIBLE], you get 400 in theta, and I'll make 10 [INAUDIBLE].

So you have flipped from this situation to this situation because now what you have is, if I have a [INAUDIBLE] here, x, then from every point, I can-- there's always theta to 400 different lenses. So this one is gone.

I can measure light coming out of this point in 400 different directions. But under each, only [INAUDIBLE]. So this is what I like. So the key lesson here is that a lens-- in the lower lens, you are basically flipping the resolution in x theta.

If you have a lens, you get more spatial resolution, but a little bit [INAUDIBLE] resolution, which is not an ideal case. The world doesn't change that much. We change what we want to [INAUDIBLE]. But in certain [INAUDIBLE] scenarios, as we see [INAUDIBLE] and so on, these aren't real. So lens flips the original [INAUDIBLE]?

AUDIENCE: My [INAUDIBLE], it depends on an object [INAUDIBLE]. But if the object was really close to the lens, it's getting a higher station in the lens.

PROFESSOR: Exactly. So--

AUDIENCE: If this location has the object [INAUDIBLE].

PROFESSOR: That's a great point. Did you hear that? I'll just try to get what he was saying. So if my point is very far away, I don't need to sample this point in 4,000 different directions, 400 different directions.

I think I can sample it in 10 different directions [INAUDIBLE]. And then having the lens [INAUDIBLE], OK? But what Rob is saying is, I can start coming closer. I'll bring the object very close, OK?

Those 400 directions are sufficiently wide, spanning 1x. And so you come really, really close, right? And it's almost 1x to 1 here. Then I really want to see all this directions.

AUDIENCE: Got you.

PROFESSOR: And then the analogy of what happens when you go from here to here is that, remember, the [INAUDIBLE] standard do a very good job of mapping the 4D light field for the real world into a 4D light field on the inside [INAUDIBLE].

I'm getting to have a little bit of analogy, like [INAUDIBLE] don't have a lens. That's not true anymore. You have a light field here and have a similar light field here for this plane. And so imagine you've got a hologram.

And the hologram has exactly 10 directions. So it's a 400-pixel hologram. It has 30 different directions. If I want to capture that, I should a lower lens here.

But if now a hologram where I wanted 10 pixels, and it has 400 different directions-- I'm so close to it-- then it makes sense to get a lower lens, so in microscopy and so on [INAUDIBLE]. So Mark [INAUDIBLE] has done some work on light fields and microscopes.

He uses different modulations [INAUDIBLE]. And the example that you saw for [INAUDIBLE] the sensor for looking at the aberration about light, there it is the main [INAUDIBLE] that are expecting to see a point that's very, very far away.

But they're not-- they aren't interested in taking an image of the point with the setup. They're just interested in finding the aberration. If they have a point very far away, and the waves are coming straight-- [INAUDIBLE] main lens. If they're coming straight, all the images are at the center.

And in the aberration, the sharp images are offset. So there are many configurations when you can decide on a main lens. And one thought experiment for you would be, what happens if there's more than one [? tensile ?] [INAUDIBLE] or more than one [INAUDIBLE] element?

So you should look up something called the [INAUDIBLE] super lens. It's a very fun concept where you actually have to put in two lenses and [INAUDIBLE] putting two lens flares right next to each other. With the right gap between them, you can create one [INAUDIBLE].

And it has very interesting properties because the focal length of the lens is in parallel with that. In a traditional lens, it [INAUDIBLE] copies of light from inside to outside inside, for [INAUDIBLE] super lens. It does a very strange transformation. Think an [INAUDIBLE].

And once you start thinking about the world as not just 2D but 4D and a photograph having not just a position coordinate but also an angle coordinate, you'll realize there are lots of other examples where this 4D representation starts making sense.

So near the end of the semester, we'll be studying medical imaging and scientific imaging using tomography and deconvolution and so on. And all those concepts in a CAT scan machine, they all work on this principle of being sensitive to position as well as angle.

So in case of a CAT scan machine, you have-- on this thing, they have a [INAUDIBLE], sometimes this chamber. And the patient goes in here. The very first class, we saw how this behaves like a [INAUDIBLE], and there's typically a [INAUDIBLE] [? engine. ?]

But what is it doing exactly? Because a set of detectors and an emitter. And your head is in here. You've done on this X-ray source, and you take this image, take shots. Then you move this light source, and you take new shots.

And this is basically capturing the light field of your body using lenses. To simplify this diagram, imagine this is light. And in case of X-rays, the source is moving, and the sensors are moving as well in a [INAUDIBLE]. But to simplify this diagram, you imagine I slap on this set of sensors, and I put the lights at different locations and just do that [INAUDIBLE].

I'm going to cast a shadow [INAUDIBLE] here, [INAUDIBLE] here. And this is basically your [INAUDIBLE], and you see an X-ray. And every ray here can be represented in the x axis.

And then it turns out that this redundancy in your space such as [INAUDIBLE] inside your body. And independent of which direction the X-ray comes in, we can [INAUDIBLE] to the same factor. So there's bone or muscle or nerve. Then, although your light field is four-dimensional, the inherent data is only two- or three-dimensional. And so you can invert that and recover the opacity on each [INAUDIBLE].

So the same thing can be done with light field. If you capture this four-dimensional X-ray, now you can go back and estimate the depth of every point in the [INAUDIBLE]. So typically, you use a stereo pair.

There's two cameras inside of some correlation to estimate correspondence and [INAUDIBLE] estimate that. But on your second assignment, what you're going to do is capture the light field and planning from the light field estimate time. And the way you're going to do that is you're basically going to line up those images.

If the point is in sharp focus in the [INAUDIBLE] be focusing, all these [INAUDIBLE] here will look the same. OK, but if something is out of focus, [INAUDIBLE] shift in [INAUDIBLE]. So you'll have a shift at some point. And then all these values will be [INAUDIBLE]. And the fact that all the values are the same indicates that now you [INAUDIBLE]. Yes.

AUDIENCE: How would we achieve [INAUDIBLE] focus? Like, focus in a [INAUDIBLE].

PROFESSOR: Yes, so that's a great point. So think about, you have a point, which is at different depths. And you're saying that in flatland, at top of image, we should focus here. In the middle of the image, we can focus here, and so on.

So this is my plane or the surface of focus, rather than x. So all you have to do is for the top point, you have to keep-- there's a little bit more you have to do. But I'll give you a very, very high level [INAUDIBLE] of how you should do this.

For this point, for this-- we have a lens here. For this particular direction, [INAUDIBLE] quite a lot of cameras. For this particular direction, you can just calibrate your [INAUDIBLE].

Put a box 1 meter wide. And I can say, within this box, at the top left, I want the box to be in the top left and merge in the back of the frame and the front. But you just calibrate that. And you can move the camera. And say that all the rays that go here, I'm going to add [INAUDIBLE].

But for the next pixel, I'm not going to add up all those rays. I'm going to add up some other rays. How do you assure that in the next [INAUDIBLE]? You know that for the point here, there'll be some value over here.

But for a point here, there is some kind of a slope. So if I just wanted to create an image that's-- so let's make [INAUDIBLE] very concrete. Before we focus on infinity, I know I should just [INAUDIBLE]. I just sum up everything along vertical lines.

If I want to put focus on my closeup, I need to shear this. So I'm going to take this, give the area same space like that. And I'm going to sum up these values. OK, and store the value. I'm going to sum up these values and store the value.

And this means I'm focusing at this point. If I want to focus on a slightly different plane, then-- so I guess vertical is center, or to focus on some other plane is a less map in focus.

So now your question is, how can I do it differently for different pixels? So on the edges, I will just do a vertical prediction. In the middle, I will do a standard prediction. And for all the places in between, I will do this line. And then I'll come back. For that, we'll get [INAUDIBLE].

And you can imagine all the crazy tricks you can do here. So one part of your segment, you're going to look at-- you're going to look at-- you could see through it. And for that, what I want you to do is take your [INAUDIBLE] cameras and then create some kind of [? a fence ?] here.

So in the real world, I'm guessing that sometimes we'll [INAUDIBLE]. It is a single part. And then there is some book or some painting back here. And you're going to take 16 images. And this is red, and this is green. You'll be able to get it [? off ?] purely by focusing on the back.

So when you focus on the foreground, you'll maybe shift a few pixels. And [? the shift ?] on the background will shift backwards. But as an extra credit, what you can do is not just refocus but also do some analysis of when you shift and add those pixels, if all those pixels have the same value or not.

So if the red object wasn't there and you just focus on the green, all the pixels here would be green. All these values would be green. But if you have some object here, some of these objects, some of these pixels would be red. Yes.

AUDIENCE: What is [INAUDIBLE]?

PROFESSOR: Here? This is x, and this is theta. Theta is the camera number, and x is still the camera lens. So the image from the first camera is placed here. The image from the second camera goes here and so on.

And because you're doing this column only one [? beam, ?] every row is independent. So think of this as only the center row of each of the system cameras. So going back just a couple of red versus green, for your first sample, you'll see, you know, refocus on green, and the object will look mostly green but in some areas.

But then you can do a simple [INAUDIBLE]. But since there's 16 cameras, and only four of them, so 12 of those are green, and only four of those are red, you don't have to sum all 16.

You can say, what's my-- what's the color that's the majority? And that color is green. So you will not get a reddish tinge to your photo, but you'll get [INAUDIBLE] red [INAUDIBLE]. If it's not a simple linear prediction, it's not just the sum of pixels. But you're going to pull out colors that are in my [INAUDIBLE].

And that will get much better [INAUDIBLE]. And again, this is in the second subpart of this assignment. You don't know how to do it, that's fine. But the main concept you want to learn is [INAUDIBLE].

When you're taking the pictures, I have plenty of instructions in the assignment. When you're taking the pictures, make sure they're placed at equal distances and you roll the camera in [INAUDIBLE] fashion. Make sure the scene, it has vibrant colors.

So this color [INAUDIBLE] is sufficiently different from [INAUDIBLE] so that [INAUDIBLE]. And before you do any of this, use the data set [INAUDIBLE] test [INAUDIBLE]. And just take 100, 100 pixels [INAUDIBLE].

And just run your code, make sure everything's fine. And then you can go and take the raw images. So you're looking for [INAUDIBLE] images. Run your code on those [INAUDIBLE], because if you start with your own images, and you're not getting correct answers, you don't know if the problem is with your photos or with your code. And does Photoshop allow you to shift [INAUDIBLE]?

AUDIENCE: [INAUDIBLE]

AUDIENCE: You can put-- you can just put a whole bunch of data as well. You don't [INAUDIBLE].

[INTERPOSING VOICES]

PROFESSOR: Go tinker that.

AUDIENCE: [INAUDIBLE]

PROFESSOR: OK, but you can [INAUDIBLE] and do the same thing. And you can just shift because you can tinker in the numbers, shift right, shift right, shift right. And you can also ask, so you can do that.

Of course, I recommend doing this [INAUDIBLE]. In MATLAB this is where you find the [INAUDIBLE], read images, shift, [INAUDIBLE]. But just because it's fine now [INAUDIBLE], don't wait till the last day [INAUDIBLE]. Start today because the fun part of this is actually taking these images and creating scenarios where you can see through.

And, you know, this should go on the Flickr page. And we will comment on each other's work. But the first assignment, only there were five or seven dozen submissions available on Flickr. So make sure your-- the first assignment was all about getting your pipeline up and running.

Almost everybody got an A, so you should feel good about yourself. But somebody got an A-plus, and others got an A. So [INAUDIBLE] good job.

So yeah, make sure it's on Flickr and all those three places, and feel free to send me an email. You can talk to Professor Oliveira or Professor [INAUDIBLE] about any of this, or also [? Ankit. ?]

And [INAUDIBLE] is also here if you want to talk about 3D scanning and so on. And use the forum on the Stellar website and so on. So today's lecture was very much focusing on the theory behind how we start thinking.

And then starting next week, we'll be looking at very different applications and different tricks we use in optics. So just go through the slides I post today. They're conceptual. But it will take you some time to just grasp it.

And again, as I said, it might feel somewhat hard to think in a dual space of x and theta, as opposed to just trying to face directly. But as you will see later, especially in your projects and so on, this will greatly simplify, this will greatly simplify the way you think about your problems. Cool? Have a good weekend.

AUDIENCE: Thank you.