MITOCW | mitmas_531f09_lec06_1.mp4

RAMESH RASKAR: The final project is a very critical part of this class. So I'll emphasize this. The assignments are simple and straightforward. If you're struggling with it, talk to me, or [? Ankit, ?] or Professor [INAUDIBLE], or Professor [? Oliveira. ?] And we are here to help you. But the final project is really, really critical.

The project has to be novel. It has to be something nobody has done before. Or at least, nobody's done it the way you are doing it. And so the problem statement has to be novel. Its execution has to be beautiful. And its impact-- you should reap some results. You should show that you are solving it or it's possible to evaluate what you have done.

And we have multiple stages to get you prepared for this final project. You have to come with at least three ideas for a final project. And this is very critical. If you go to the wiki on the camera culture group page, we have a whole section on how to come up with great ideas and how to brainstorm and so on. So I encourage you to do that. On our [? Stellar ?] page we also have a whole presentation on how to come up with new ideas. There are six ways of coming up with new ideas.

And you should start setting up a meeting with me or any one of these people in the next week. If you're in the building, it's easy. You can just catch me. Usually, between 5:00 PM to 6:00 PM is the best because there are no officially scheduled meetings. You can just come to my office, and we can chat.

If you're not taking this class for credit and you're a listener, I would really appreciate if you will pitch an idea for a project. And we'll have an opportunity to do that on October 30. So just come up here and say, I have this crazy idea. If somebody wants to work on this, let's team up.

And then I want you to send a very simple email with-- after discussing these three ideas, I, or [? Ankit, ?] or our other two mentors will help you narrow down to maybe the top two ideas, or maybe the top one idea. And then, for that, you need to send these five things. All right?

And on November 6, we will do a three-minute presentation so everybody knows, in the class, the kind of problem you are attacking. Maybe there's some synergy between multiple projects. You can help each other find software, or talk to people, or get some equipment. And then final proposals are due two weeks before the actual final presentation. And by then you should have some initial experiments.

Remember, your final assignment is due on November 13. So you can't start thinking about your final project after the final assignment is due. You need to start thinking now.

And at this stage, it doesn't have to be completely hashed out. That's why you're in the class. We'll help you think through that. And then December 4, the class is finished. You finish really early, because as I said, we meet on Fridays. And the Thursday after this, so December 10, is the last day of classes. So we are one of the first classes to finish.

Unfortunately, November 27 is a Thanksgiving break. And so-- which means we don't have a class just before our final projects to discuss and do other things. But I will be available throughout that time. And we can help you. If you need equipment, you need cameras, you need software, whatever-- we can try to help you. Any questions on final projects?

And Mike was saying that it might be good to know what students did last year in-- as last projects, and how they decided, and so on. So Mike Hirsch is going to come and talk today about his [? pilot ?] screen. And that would be a good way to ask him, how did he get started, and how did they consider the paper?

So today we're going to talk about lightfields-- just finish up the lightfields, and then talk about cameras for SCI. So we said there are basically three ways of coming-- of capturing a lightfield. Can somebody tell me in two sentences what's a lightfield?

AUDIENCE: On the real cameras?

RAMESH RASKAR: In a real-- inside a camera? What is it? How are they represented mathematically? A representation-- just a representation.

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: So it's-- how many dimensions does it have [INAUDIBLE]?

AUDIENCE: Four dimensions. Four.

RAMESH RASKAR: It's four-dimensional. Although the world seems to be two-dimensional, [INAUDIBLE] actually four-dimensional. And where do we get those four dimensions from? We have a lens. And we got a sensor. So in flatland it's just two-dimensional [INAUDIBLE] wall. Where are the two dimensions coming from? Well, the x dimension is easy. That's one dimension. And the other dimension is?

AUDIENCE: Through [INAUDIBLE].

RAMESH RASKAR: [INAUDIBLE], That's this dimension. So if you connect any two points, that indicates the direction of the rate. And then flatline, [INAUDIBLE], all the raised bits is two dimensional. In the real world [INAUDIBLE] it's four dimensional.

Because your sensor is going to be X and Y. And your lens is going to have [? VW. ?] So that's the lightfield that we care about. And why do we care about lightfields so much? What's so unique about lightfield?

AUDIENCE: If we capture the lightfield, we have captured almost everything that we can capture.

RAMESH RASKAR: Exactly. It's a complete representation of light that's entering the lens. So anything that you could ever imagine doing with dimension-- changing focus, changing zoom, changing aperture size-- all those things are already captured in this four [INAUDIBLE]. And from there you can do anything you want.

And if you cannot do it using mechanically changing camera parameters, you cannot do it with lightfields. Sorry-- if you cannot do it with lightfield, you cannot do it by changing [INAUDIBLE]. So it's a very, very powerful way of-- and as we saw earlier, the lightfield also represents the [? waveguard, ?] both phase and function of [INAUDIBLE].

So given that this is so important, photographers never thought about capturing the lightfield. [INAUDIBLE] capturing light. And that's why this is computational-- from a pure computational camera, because we're trying to understand this relationship. And this is what allows us to build cool camera toys, and also come up with [INAUDIBLE].

So given that, what are the three ways we can capture this? One is using a lens [INAUDIBLE] that we saw earlier [INAUDIBLE]. All you want to do is, right now, at this pixel, a ray from here and a ray from here. They converge. And any variation in the radius along those rays is lost. We want to make sure that we can capture this and this individually.

So what are some ways we can do it? Instead of putting the sensor here we can put the sensor further back, and then [INAUDIBLE]. And the light can [INAUDIBLE] here [INAUDIBLE] for the lighting. [? Inputs ?] here. And it'll be mapped [INAUDIBLE] different pixels.

So you have [INAUDIBLE] equals 0, [INAUDIBLE] equals this [INAUDIBLE] and plus 4, and [INAUDIBLE] minus 4. Then have captured-- for this given x, I've captured-- let's say this is 1, 2, 3, 4, 5. Then have captured their 3 comma-- well, minus 4.

AUDIENCE: Minus 4.

RAMESH RASKAR: And yet I've captured 3 comma [INAUDIBLE]. So I have captured all [INAUDIBLE]. But of course, you won't have what?

AUDIENCE: Resolution.

RAMESH RASKAR: You won't have resolution. Because for x equals 3.5 they're just opaque. So that light is lost. [INAUDIBLE] captured.

So if I say here-- let's say I have 1,000 pixels. So that's 900 pixels [INAUDIBLE]. And I'm going to chop it into [INAUDIBLE] plus 4, minus 4. What is the final-- how many pinholes can I put here? So this total resolution is x resolution times [INAUDIBLE].

In a traditional camera, 900 is just your execution. But now we're going to also try to capture the [INAUDIBLE] variations in the same number of [INAUDIBLE]. So we know that our [INAUDIBLE] resolution is 9. So [INAUDIBLE] 100. Which means that our image is only going to be 100 pixels.

[INAUDIBLE] only 100 pinholes. And for each pinhole I'm going to get the image [INAUDIBLE] pixels. And from this 100 times 9 image, which is 900 pixels, I can create a lightfield where the spatial resolution is 100 and the angular resolution is 9.

And what's the disadvantage? One is that you lose resolution.

[INTERPOSING VOICES]

AUDIENCE: Light.

RAMESH RASKAR: You lose light, because all the lights that's going through this opaque area is completely lost. It's like looking through multiple pinhole cameras. It's like looking at the world through [INAUDIBLE] holes. So most of the light is lost.

In this case, let's say out of nine pixels, one pixel is open. Then only 1/9 of the light is being captured. The other 8/9 of the light is lost. So that's a big problem.

But still, conceptually, this is very key, because you can say, I want to capture space-- space radiation as well as angle resolution-- angle radiation. So I'm just going to chop my [? word ?] into 100 pinholes and 9 angle-- [INAUDIBLE] angle spaces.

And I'm going to capture nine pixels, and so on. It's a very clean, simple model. But in the real world, this is very inefficient. It's just like for a camera, if you're [INAUDIBLE] pinhole model-- pinhole camera model. But in the real world, [INAUDIBLE]. Now--

AUDIENCE: So let me ask you this. Can you really construct this in such a way that, as you go to the end of the-- I mean, your x. Aren't you wasting part of the nine spatial samples that you have?

RAMESH RASKAR: You mean at the top or the bottom?

AUDIENCE: Yeah. Because maybe you can't actually have this [INAUDIBLE] opening over the entire set. Yeah, exactly.

RAMESH RASKAR: As long as you-- this is your sensor. And you can create a pinhole that can capture light from different directions. You're OK. There's not much you can do.

I think you might-- what you may be thinking about is, this pinhole, at some angle, will become really opaque. You won't be able to see through it. Is that what you're saying?

AUDIENCE: Yeah. What I'm saying is that maybe, through this particular angle, the pinhole doesn't span the whole nine pixels you have behind it.

RAMESH RASKAR: That's a good point. But you are leading me to the next question, which is, this spacing here. What is magical about this spacing? Because if you don't have the right spacing, you'll get those problems. What needs to happen in the spacing so that you actually capture all these 900 [INAUDIBLE]? Yes?

AUDIENCE: [INAUDIBLE] they really should not overlap [INAUDIBLE].

RAMESH RASKAR: Exactly. Exactly. So if you think about the blob that's coming through from top to bottom, there'll be some loss here.

If you see the blob that's coming through here, these blobs are just barely touching each other. So here we have theta equals plus 4, 3, 0, 1, minus 4. And then the next one is theta equals plus 4. So it's flipped [INAUDIBLE]. Yes. Then, from minus 44, you guys want to go from minus 44 [INAUDIBLE]

Now, if I move this pinhole further or back, you're going to get [INAUDIBLE] variables, right? Let's say I move this pinhole a little forward and I just take one pinhole [INAUDIBLE]. And everything else is the same.

Are the blobs larger? If I place the pinhole at the same-- I place the pinhole exactly at 9 pixels away. If I can put it exactly 9 pixels away, the rays on the [? cloth ?] will slide into [INAUDIBLE]. So this blob here will start interfering with this blob here [INAUDIBLE]. So that's one issue. What happens if it's too close? Instead of moving further away, it gets too close?

AUDIENCE: No spaces.

RAMESH RASKAR: There will be a space between those blobs. Is that here [INAUDIBLE] a blob here and we get a blob here. And some pixels will be missing. So this is a very important point. And this is related to the numerical aperture, or the F-stop of the lens.

So an F-stop of the lens-- again, pardon all this terminology. I don't like it. But [INAUDIBLE] about it is the f-number is simply the ratio of the diameter of the lens with the distance to the sensor. So it's a diameter. And your distance of the [INAUDIBLE].

The relationship between the two is just c over [INAUDIBLE] so if you have a-- and remember, the diameter is in the denominator, which means that as you have a larger lens, the f-number actually goes down. [INAUDIBLE].

So let's take a concrete example. Let's say my lens has a diameter of 25 millimeters. And my focal length and distance is about [INAUDIBLE] millimeters, then the f-number is what? 50 [INAUDIBLE]. Does anybody know what the f-number of [INAUDIBLE]? [INAUDIBLE] 0218. Then there is 4 and so on.

So as you can see that if I make my diameter half-- so instead of 25 millimeter, I have 12.5 millimeter-- then what do I get here? I get 50 divided by 12.5. So that's [INAUDIBLE] the f-number of a 12.5 millimeter lens with a 50 millimeter focal length is 4.

From 2 we jump to 4. What should we do to jump from 2 to 2.8? This is where the [INAUDIBLE]. But maybe some of you have forgotten [INAUDIBLE]. Yes?

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: [INAUDIBLE]. How do you go from 2 to 2.0 [INAUDIBLE]. But how do you get [? 4.4? ?]

AUDIENCE: [INAUDIBLE].

RAMESH RASKAR: So what's [INAUDIBLE]?

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: 1.41428 something, right? As far as [INAUDIBLE] are concerned, it's [INAUDIBLE] because we like to think about this size numbers. And this problem [INAUDIBLE] even bigger problem once you go for [INAUDIBLE].

Anyway, so why-- where are we getting this [INAUDIBLE]? If you go from [INAUDIBLE] when you go from 25 to 12.5, the amount of light that's collected by this [INAUDIBLE] across this distance is related by what factor?

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: 5 times 4. So the diameter is [INAUDIBLE] plus 2, but the area is decreasing by a factor of 4.

Now, if you want-- so four times [INAUDIBLE] coming in. On the other hand, if you want to go from 25 to some other diameter so that you get half the width [INAUDIBLE]-- if you divide 25 [INAUDIBLE], whatever that is-- 18, something like that?

25 [INAUDIBLE]. Let's say it's 18 point something. That fair? [INAUDIBLE]. I'm being imprecise. [INAUDIBLE].

So when I go from here to here, [INAUDIBLE] of 2 in diameter, I get half the length. And that's what photographers mean to go by-- go down by 1 F-stop. So when you go down by one F-stop, are you going from 2 to 3? Are you going from 2 to 4? What are you doing? It's such an imprecise amount.

When you're going by F1 F-stop, it means going from 2 to 2.8. It's completely unnecessary [INAUDIBLE]. So just get rid of this terminology altogether. And if you really want to think about-- the reason why it's worthwhile in thinking about the ratio of diameter to the focal length is about the amount of light.

And can I tell you why that's the case? So let's say I have a 25-millimeter lens with a 2.5-- sorry, 50-millimeter focal length. Or let's say I have a 5-millimeter lens with a 10-millimeter focal length. This system and this system have the same exact f-number, which is 2. Here it's 50 by 25. Here it's 10 by 5. So both of them have the same f-number.

What's constant between the two? This angle here. This angle is the same in both systems. And this is critical, because when you are looking at the [? world, ?] you want to say, at the given pixel, over what [INAUDIBLE] am I [INAUDIBLE]?

And so that's a very important factor when you think about how many photons you want to capture, because the angle-- the [INAUDIBLE] angle [INAUDIBLE] from the light determines how [INAUDIBLE], and certainly helps you to think in terms of ratios of angles. Other than that, [INAUDIBLE] terminologies [INAUDIBLE]. What happens after 4? What's the next f-number?

AUDIENCE: 5.6.

RAMESH RASKAR: 5.6. What do you think 4 times square root of 2 is? Slower [INAUDIBLE].

AUDIENCE: 5.64.

RAMESH RASKAR: It's probably 1.4 times 4. So it's 4 point-- yeah.

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: 5.64, [INAUDIBLE]. Getting more and more imprecise. And as you can imagine, as you go further away, you-- so this is the same confusion between whether a megabyte is 100 kilobytes or 124. Actually, it's even more confusing than this.

And I heard a very interesting story at [INAUDIBLE]. When the economy was not going so well, in the contract they said, we will charge you this much for each megabyte or each gigabyte. And [INAUDIBLE] a kilobyte-- the difference is very small. It's less than 2%. [INAUDIBLE] larger gigabyte to terabyte. Difference between the power of 10 and power of 2 is actually significant.

So this woman wakes up one day and says, you know what, I'm gonna charge people by the bytes-- the [INAUDIBLE] power, not [INAUDIBLE] power. And so they [INAUDIBLE] by [INAUDIBLE].

And so the same kind of problem is creeping in here. [INAUDIBLE] square root of 2 is 1.4. It starts getting more and more confusing as you go. And then, as you know very well, after 5.6 you start jumping. You have 8, and 11, and 22, and [INAUDIBLE]. And then a [INAUDIBLE] plenty more. It just gets too confusing. So just ignore [INAUDIBLE]. It's the most confusing thing. Way down by F-stop, I think I should go ahead and approve [INAUDIBLE].

Photography school [INAUDIBLE]. So what we're trying to do here is our angle here is decided by this ratio, [INAUDIBLE] to focal length. And the angle here is this-- again, a ratio of the blob to this distance. So let's call this C1. And let's call this [INAUDIBLE].

So what you want is the ratio of e-- sorry c to d to be equal to the ratio of l to your [? losses. ?] And if that is matched, then all your blobs are going to just barely touch each other. If it's not matched, then either you'll get [INAUDIBLE] or you'll get [INAUDIBLE]. That's the basic math behind a light frame.

So it's very conceptually, but as we know, it's going to block the light. And [INAUDIBLE] even talk about [INAUDIBLE]. Has anybody done, here, pinhole photography? So what are the problems you've faced in your pinhole photography?

AUDIENCE: It's just, there's so little light [INAUDIBLE]--

RAMESH RASKAR: So little light, so--

AUDIENCE: --exposure time.

RAMESH RASKAR: Exposure time is very long. And? And the image quality?

AUDIENCE: It's blurry, because you could never [INAUDIBLE] the small holes.

RAMESH RASKAR: Exactly. The image is blurred because of refraction. So if you're viewing in a pinhole, then light comes in. And it actually doesn't go in a straight line. It actually bends a little bit. And because of that, the single point in the [? world ?] doesn't map to a single point in the sensor. But it maps to a small [? blur. ?]

And this is very similar to the analogy of just using a water hose. If you have a water hose, and it's open, water comes in. And you get always the same thickness, same width of the water flow, right? But as you start shifting this, eventually the water just starts spraying [INAUDIBLE].

When the size of your opening in the water hose becomes comparable to the molecules of water, it actually starts spraying out. And that's diffraction, as well. And we have the same principal here. In very simple words, we have photons coming in here with certain wavelengths. What's the wavelength of green light?

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: 500 nanometers. Remember these numbers. They're very, very [INAUDIBLE]. So 400 to 700 is blue, green-- and this is nanometers [INAUDIBLE].

So if the light is 500 nanometers, which is 0.5 micrometers, and your pinhole starts [INAUDIBLE] about 1 millimeter, which is 1,000 micrometers, then you're probably OK. But as you start weighing millimeter and below-- which is what your cell phone cameras are, fortunately-- you start getting to 500 microns, which is half a millimeter. Then your wavelength is comparable to the size of your photon. And you start getting this [INAUDIBLE].

So the relationship between the size of opening and the size of the wavelength kind of decides the diffraction. And the focal length is very easy. This one here is simply-- [INAUDIBLE] I said the aperture is 8.

This angle here is going to be [INAUDIBLE] in gradients [INAUDIBLE]. We won't be talking about it too much in this class-- in this lecture. [INAUDIBLE] an indication of how quickly the light is turning up.

So in the worst case scenario, when you have the big hole the same size as the wavelength of light-- so let's say you created a pinhole whose width equals [INAUDIBLE] nanometer, or 0.5 micrometer, what [INAUDIBLE]? If you're wondering, 1 gradient is how many of this, [INAUDIBLE].

So this cone will be 57 degrees. So it's part of the question. Even if you can just do calculations, and even if you plot something that's 1,001 millimeter-- so that's [INAUDIBLE] micrometers-- the angle's [INAUDIBLE] pretty wide. That's what you need to know.

And so when camera makers are selling you lenses-- really crappy lenses on a mobile phone camera, but just giving you a [INAUDIBLE] resolution of 5 megapixels and 10 megapixels, it doesn't make sense because your image will be blurred. You don't need that many. So other way to think about that is, can your camera capture all this?

AUDIENCE: Mm-hmm. [INAUDIBLE].

RAMESH RASKAR: You should just take photos later on. If your cell phone camera has a typical aperture size of, say, 2 millimeters, that's [INAUDIBLE] and then your sensor light comes in, you're kind of zooming in [INAUDIBLE]. And it kind of spreads out. And let's say it spreads out over only about 20 micrometers.

The pixel of a camera is about 5 micrometers each pixel. So if the blur is already 20 micrometer, it doesn't make sense to have a pixel that's that small. You should just have pixel that match. [INAUDIBLE].

If someone's selling you a 10-megapixel camera with a 5-micron pitch, they should just give you something that's much lower resolution than that, because the numbers won't make any sense. But again, this is a typical marketing gimmick camera makers will use, where they will sell really high resolution sensors, although clearly, you cannot really capture that big [INAUDIBLE] resolution. Skip.

And there's a new trend. I think camera makers have started selling [INAUDIBLE] megapixel [INAUDIBLE]. And now they have to start selling cameras with lower resolution. I think the latest Canon has lower resolution than its previous version. I believe the G6-- I think it's 12 megapixel. But at least now it's dynamic. Because they realized it doesn't make sense to just keep boosting the megapixels when the aperture of the lens is [INAUDIBLE].

So recently, I was working on a project where they wanted to capture 50 gigapixels. 50 gigapixels. 50 times 10 to the 9. This is usually 2 times [? 10 to the ?] 6 pixels. And as the megapixels increased, the size of the-- as you can see, this ratio here-- as the megapixels increased, the size of the lens required increases correspondingly. And that's why expensive cameras have much bigger lenses.

AUDIENCE: [INAUDIBLE]. I had a question. [INAUDIBLE], wouldn't it be much more effective to come up with a measure [INAUDIBLE] size of the camera [INAUDIBLE]? [INAUDIBLE] can't really [INAUDIBLE].

RAMESH RASKAR: Exactly.

AUDIENCE: [INAUDIBLE].

RAMESH RASKAR: And the camera makers [INAUDIBLE].

AUDIENCE: Yes. I mean, of course it's a marketing thing. [INAUDIBLE].

RAMESH RASKAR: Yeah, yeah.

AUDIENCE: [INAUDIBLE]?

RAMESH RASKAR: Yes, there is. So there are all these measures called modulation transfer function, and space-bandwidth product, and so on. And if you want, we can discuss about that later on. It's not so simple [INAUDIBLE] what we're doing here.

But [INAUDIBLE], yes, let's get more resolution, or, this [INAUDIBLE] resolution. Those are the laws of physics. You can have a lot of holes, but [INAUDIBLE]. So the other way to think about this is, how can you create a [INAUDIBLE] that creates a spot that's [INAUDIBLE] start decreasing your [INAUDIBLE] size for the things [INAUDIBLE].

And it's almost-- it's funny how nature works. There are always, always limits as we get closer to certain [? questions. ?] If you think about the density of water, it decreases as [INAUDIBLE] temperature. Or, at some point, [INAUDIBLE] potentially starts [INAUDIBLE] diffraction, or [INAUDIBLE]. [INAUDIBLE] where it seems like [INAUDIBLE] puts some stuff [INAUDIBLE].

So we're talking about resolution. We're talking about the ratio of the diameter to the length. There's a very minor and very subtle point here. I'll just make it for those of you who are thinking about [INAUDIBLE] focal length, effective f-number or [INAUDIBLE] number. Which is-- although we're defining it by the diameter to the focal length, the sensor is never [INAUDIBLE], because that's really the case when you're looking at [INAUDIBLE].

When you're looking at something closer, the sensor is-- say the focal length [INAUDIBLE] millimeter, the sensor is 50-plus, like 51 or 52, maybe. But again, that's small enough that [INAUDIBLE] a lot of times, where [INAUDIBLE] like that. It's by taking [INAUDIBLE].

So we're talking about pinhole. Then we started talking about how pinholes aren't great because they lose light. Exposure ends up too long and [INAUDIBLE]. And all those problems are going to appear here as well. It's just that what we have created here is an array of 100 virtual cameras.

So 100 virtual cameras where each camera has a pixel-- a resolution of [INAUDIBLE] pixels. [INAUDIBLE] one big camera into 100 cameras, each with [? 9 ?] pixels. And in the real world, it would be 10 [INAUDIBLE] cameras, [? 100/100, ?] each with a resolution of [? 9 ?] megapixels. That's what makes it interesting. From one camera [INAUDIBLE] 100,000 angles. And as we saw last time, you can do it at home. [INAUDIBLE].

So just like in regular cameras, we don't use pinholes. We use lenses. You want to replace each of these pinholes, now, with lenses. And that's why we call it [INAUDIBLE]. So I'm just going to draw on top of this with a different color. Tell me if it gets too complicated. I'm just going to draw the landscape.

And it's going to do almost the same task. We know that at the center of the lens, [INAUDIBLE]. Because remember, a lens is made up of [INAUDIBLE]. And the middle one is just a sheet of glass. So when it goes to the center, nothing changes. [INAUDIBLE].

When you use a lens [INAUDIBLE], we can do the same thing. But there's one very special-- two, actually, very special constraints we have to achieve to make this happen. Because it's a lens and we're trying to form an image.

When we're converting light to the camera, what we're really doing is forming an image of the lens on the sensor. You remember, we heard this concept of imaginary plane, a point in 3D as well as pointing a location on the-- sensor position is what [INAUDIBLE] the picture, the corresponding [INAUDIBLE].

And here, what you want to do is create a lens which forms an image of the lens all [? the same size. ?] [INAUDIBLE] this focal length and its distances in such a way that if I put a point here, I get a sharp image of the point here. Because in other words, [INAUDIBLE]. And each detail here has to map to 1 pixel exactly.

So that makes it extremely difficult. Because now we're almost in a microscopic point. And [? I do ?] create these lenses extremely high quality so that an image of a point can-- at about 50 millimeters is formed at some very small distance, usually 1 [INAUDIBLE] or 1 millimeter. It's about 500 microns [INAUDIBLE].

So this becomes very challenging. This is a very specific constraint. There's exactly one plane in which I can put the sensor. If I put it too far here, then the image of the lens on the sensor will be blurry. If I put it too close, again, the image will be blurry. And that very special constraint makes building a lightfield camera out of lens [INAUDIBLE] is extremely challenging and very expensive. [INAUDIBLE].

So let's say you-- by the way, all these things we talk about where if you put the lens stack-- you can put the lens a little bit further out and change its focal length so that this one is still imaged here. But conceptually, it's to the pinhole.

So if you put the lenses over here and the ratios of this to this is not matched to this to this, then you'll get overlap. And if you put it too close and change the focal length accordingly, you'll get dark spots [INAUDIBLE]. So we still have to worry about a standard set of issues.

So this additional thing you have to do, which won't look-- a lot of people forget-- which is, if there's some points-- there's some point here out in the world [INAUDIBLE] again [INAUDIBLE].

And what we're trying to do is, this point is in focus over here. And then we're going to chop this into nine [INAUDIBLE], and [INAUDIBLE] plus 4, [INAUDIBLE] minus 4. And the one that goes to the center gets divided too plus [INAUDIBLE]. That is very focused. And what we have done is we have captured the lightstream of [INAUDIBLE].

If I want to create a lightfield, I [INAUDIBLE] differently over here. The question is, can we-- what's happened with-- remember, we made this claim that if I capture this 4D lightfield, I can do anything I want. I can focus here. I can focus there. I can use the light. I can change the aperture size. I can change the focal length. I can do anything I want if I capture that [INAUDIBLE].

So one question to be thinking about when you are doing this assignment is [INAUDIBLE] capture the light for one particular plane, how do you recreate it [INAUDIBLE], either by refocusing it or [INAUDIBLE]? What's the difference between this situation and the assignment that you're doing?

Because [INAUDIBLE] cameras, and maybe last time we realized that the lens is what? It's the area of impulse with [INAUDIBLE]. I can chop the lens as [INAUDIBLE]. You can take off the lens and you can chop it up. And you can treat it the same as these pinholes, which corresponding to something next to it.

In this case, just a sheet of glass. It's a pinhole plus a prism. When you take a picture with a camera array, unfortunately, you're going to have to do that computation. All you have is a set of cameras.

AUDIENCE: [INAUDIBLE].

RAMESH RASKAR: Thanks. All you have is a set of cameras. You don't have this prism. And mathematically, you're going to shift the image, which is equivalent [INAUDIBLE]. So you're going to take [INAUDIBLE] images and you're going to shift them by plus 2 pixels, plus 1 pixel, 0 pixel, minus 1 pixel, minus 2 pixel. And that's the same as putting a set of prisms. Is this analogy clear?

So once we have done all this, there's two classic ways of thinking about capturing lightfields. Again, 1908 is when this idea started coming up, [INAUDIBLE]. But the practical solutions came much later. Remember, back then, it [INAUDIBLE].

Meanwhile, a third solution came on just two years ago, whilst capturing lightfields. And we're going to look at that for the next few minutes and then switch over to [INAUDIBLE]. Anybody has thoughts on some other ways of capturing the light [INAUDIBLE]? Yes?

AUDIENCE: [INAUDIBLE].

RAMESH RASKAR: Excellent. Excellent. So you [INAUDIBLE] lots of [? hats ?] to go around this. So one very interesting type these guys came up with was they say, I don't want to put anything close to the sensor-- just a traditional camera.

And [INAUDIBLE] this. I'm going to do this part up here. Because all I care about is this pixel, I want to know how each of these rays [INAUDIBLE] what is the radius of each [INAUDIBLE]? So all I will do is I will block this part of the lens. [INAUDIBLE] and that will give you this direction for each pixel.

Then I take the second photo, but I will block everything except this region. [? 84. ?] And if I take nine such photos, then [INAUDIBLE] getting 100 times 9, what will I get? [INAUDIBLE] each has a resolution of 900 [INAUDIBLE].

So we get 900 [INAUDIBLE] this whole time. And this is a single [INAUDIBLE] and 9 times [INAUDIBLE]. So this is what [INAUDIBLE]. And they came up with better [INAUDIBLE], but better than this class.

But you still learn it's not a new matter, because it's using the same basic concept that [INAUDIBLE]. You have something of your own?

AUDIENCE: [INAUDIBLE].

RAMESH RASKAR: Yeah, OK. What else? Think about all the-- that document online, that presentation about how to come up with new ideas. One way to think about it-- think about all the cameras you know, all the [INAUDIBLE]. On a camera, you have [INAUDIBLE] exposure time, focus, focal length, moving the camera, wavelength.

Think about all the ways you can use those parameters to reach the scope. [INAUDIBLE]. This is called x [INAUDIBLE]. And you know what it means. Yeah?

AUDIENCE: What if you got a-- so you're shooting out of the [INAUDIBLE].

RAMESH RASKAR: Why does it have to be a thing?

AUDIENCE: Yeah.

RAMESH RASKAR: [INAUDIBLE] something else.

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: Exactly. What [INAUDIBLE] should it be?

AUDIENCE: [INAUDIBLE].

RAMESH RASKAR: Like on some kind of a concave or [INAUDIBLE] sphere.

AUDIENCE: A similar [INAUDIBLE].

RAMESH RASKAR: Remember, CAT scan machine works on the most simplest [INAUDIBLE]. And a CAT scan machine, it is [INAUDIBLE]. [INAUDIBLE], Or even thinking of doing some non-interfering optical communication.

[INAUDIBLE] interested in this. You will do the same thing. [INAUDIBLE] are coming from different directions [INAUDIBLE] them, as we told them [INAUDIBLE]. We'll put the area of angles [INAUDIBLE]. [INAUDIBLE] a bigger project.

AUDIENCE: [INAUDIBLE].

RAMESH RASKAR: So let's think about this third solution, which is a relatively simple idea to explain, actually. What we're going to do is not use lens [INAUDIBLE], but use a [INAUDIBLE]. I'll show you how this works, and then I'll explain how it works.

All I'm going to do is place a mask-- instead of a pinhole array, we're going to place the mask that has certain consequences. So here, it's a pinhole array. They're going to be, basically, some [INAUDIBLE]. It's going to have some strange effect.

You can bring this [INAUDIBLE] and put it [INAUDIBLE]. And if you do that, it turns out, you'll get pictures that look like this. And if you try to force it out of focus, you have this really strange [INAUDIBLE]. Part [INAUDIBLE] focus, it looks like light has been [INAUDIBLE] a little bit. But they're the same I think. [INAUDIBLE]. But for out of focus, it's this really strange [INAUDIBLE]. And I'll summarize how it's computed and then come back and explain why it works.

And all we're going to do is take a traditional camera. [INAUDIBLE]. You could [? take ?] any image, in general, and just take its Fourier transform, which is what JPEG compression will do as a [INAUDIBLE] step. You realize that most of the energy in this Fourier transform is in the low frequencies including you put low frequencies in the center. One day you will [INAUDIBLE] center [INAUDIBLE].

So most of the energy is in low frequencies, low spatial frequencies. So in the center, you have the DC component, which is the average of all pixels. And then you have first frequency, which is how many [INAUDIBLE] you can put in a dimension, and so on.

Most of the [INAUDIBLE] over here, if you place this very high frequency mask where the pinhole layer was, and then [INAUDIBLE] Fourier transformed, it looks really, really strange. It has [INAUDIBLE] energy in high frequencies, as well.

And those of you who are used to looking at oscilloscope or radio frequencies and so on-- if you just capture the radiowaves, all the radio stations, and look at the spectrum, it will also look something like this. It will have some carriers in the middle. And those will get [INAUDIBLE] around the carrier. And there will be a [INAUDIBLE]. So a 99 megahertz station is transmitting its audio over [INAUDIBLE] stations and transmitting the audio [INAUDIBLE] and so forth. So this is what the spectrum will look like if you just capture any kind of [INAUDIBLE] signal.

And something similar is happening here. And I'll come back to this and explain how this example works. So that-- and what we're going to do eventually is take the truly Fourier transform, which looks like this, and we're going to shift. We're going to take this 2D wall and make it into a 4D hypercube.

And we're simply going to take this inverse Fourier transform and recover all those images that you will have captured if you placed the camera at different positions. It's as if we have taken the lens and split it into 81 different cameras. [INAUDIBLE].

As you can imagine, from these 81 cameras we have captured-- so for every one of those slices on the lens, we have captured an image that's only 200 pixels by 200 pixels. But we have created 81 such pictures. And so here you are just seeing some of those different pictures, a few [INAUDIBLE] pictures.

But using this 81-camera array-- the box I showed you last time was a 5 by 5 camera array. And you can think of this as an 81-camera array. But we didn't build a new device. It's just an ordinary SLR camera we didn't format. And we just placed the smart sensor. And certainly, we have-- for two extra dollars, we have now 81 cameras. But of course, each of them is way more efficient.

AUDIENCE: What's the advantage of this mask to the pinhole-- or pinhole mask?

RAMESH RASKAR: So the question is, what's the benefit of this type of mask over a pinhole array, or some [INAUDIBLE]? What are the disadvantages of pinhole arrays?

AUDIENCE: More light is going to [INAUDIBLE].

RAMESH RASKAR: Right. So it's more light, because the-- almost 50% of the light will go through. So previously, we know the pinhole light is-- very little light goes through. What's the second problem [INAUDIBLE]? Yeah?

AUDIENCE: Because the image you get through a pinhole [INAUDIBLE] is not established [INAUDIBLE].

RAMESH RASKAR: Exactly. If you [INAUDIBLE] diffraction arrays. Even here, you have some diffraction [INAUDIBLE] above the [INAUDIBLE].

AUDIENCE: So it's basically assimilating. And that's not very--

RAMESH RASKAR: Exactly.

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: Well, [INAUDIBLE]. What you capture looks like almost not our image, except those very strange effects are called [INAUDIBLE].

AUDIENCE: So but are you saying you're applying the inverse transform to the small sessions of this spectrum, right?

RAMESH RASKAR: No, not really. Not really.

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: We'll take this full 2D image. And we're going to [INAUDIBLE] it and create a 4D hypercube. So let's make it simple. Let's say the image was only 1D, just this part of the line here. If you take this frequency transform-- actually, [INAUDIBLE] transform looks like that. If you take this frequency transform, you will capture-- there will be some 1D frequency transform that looks like this.

And what we're going to do is-- so let's go step by step. So the frequency domain-- this is a little bit of a [INAUDIBLE]. We'll come back and talk about 4D techniques in a couple of lectures. And this will become more clear at that time.

But basically, we have some variation in x and some variation [INAUDIBLE]. We saw there. And that's because it was 100 by 9. In this case, it's 200 by 9. That's 200 [INAUDIBLE] actually. 200 by 9, we're just thinking [INAUDIBLE] data.

And then our sensor, however, is just 1D in this case. So although we want to capture 200 here and nine here, we cannot capture that much from the sensor. So what we basically do is take that 1D sensor-- [INAUDIBLE]. We're going to take that 1D sensor, and it's going to capture different parts of the signal. And then we're going to reshape it.

We're going to chop-- we're going to chop this part over to here, chop this part over to here. This one stays in the same place. Take this part [INAUDIBLE] here. And this part, goes over here. And now, from a 1D signal, you have created a 2D length. And it's not really visually clear, I'm sorry, because [INAUDIBLE] scheme. [INAUDIBLE]. And from that, we can recover--

AUDIENCE: So you have all the coefficients for the entire image.

RAMESH RASKAR: Exactly.

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: The whole coefficient [INAUDIBLE]. Yes?

AUDIENCE: What about the [INAUDIBLE]?

RAMESH RASKAR: Sorry?

AUDIENCE: The [? reason. ?]

RAMESH RASKAR: Frequency domain [INAUDIBLE]?

AUDIENCE: Yes.

RAMESH RASKAR: So we'll talk about that. If you think about this particular problem where we have overlap or undershoot, there's also a [INAUDIBLE]. Because we're trying to capture more signal, but we don't have enough bandwidth here. And the same exact problem here as well. So you;re going to assume that the scene doesn't have [INAUDIBLE]. So we're [INAUDIBLE] and must [INAUDIBLE].

AUDIENCE: But doesn't it mean that you lose [INAUDIBLE] data [INAUDIBLE] equipment?

RAMESH RASKAR: You can only-- so you [INAUDIBLE] only 200 pixels in the world.

AUDIENCE: Even if the spectrum is not [INAUDIBLE], it's very simple. You need to sample [INAUDIBLE].

RAMESH RASKAR: You have to hope that-- you hope that your spectrum looks something like this. If your spectrum actually has lots of energy up here and down there, then, yes, you'd get [? RPS. ?]

AUDIENCE: Even-- and even you can separate this signal with a much lower frequency than the spectrum [? specific ?] right?

RAMESH RASKAR: Repeat that?

AUDIENCE: You can separate in much lower frequency than actually signal separate?

RAMESH RASKAR: Using what? Using how much [INAUDIBLE]?

AUDIENCE: Use the tool. No, just simply low frequency.

RAMESH RASKAR: Just take--

AUDIENCE: For this center one.

RAMESH RASKAR: Yeah. That's exactly what happens in JPEG. You take whatever 4-megapixel image, and then you're able to represent it as a half-megapixel image, because you're able to take only the lower frequencies-- lower frequencies and represent them above. And you take the higher frequencies, but we don't represent them [INAUDIBLE]. That's how JPEG [INAUDIBLE] frequency [INAUDIBLE].

AUDIENCE: So you assume--

RAMESH RASKAR: But we cannot do that for real. We can do that in optics.

AUDIENCE: Yes.

RAMESH RASKAR: You can do it in software. So here's a method to do that-- not the [INAUDIBLE], but the remapping in optics.

AUDIENCE: So you still assume it's overcycle?

RAMESH RASKAR: Yeah, exactly. So if the world actually had a checkerboard, an extremely high resolution checkerboard, then you'd have problems. It's a similar effect-- your software is where you really find structure. And you just take the picture. You see these areas and artifacts.

Or, if you have a-- if you take a fence, we have [INAUDIBLE] take picture [INAUDIBLE]. So it's the same problem. It's just [INAUDIBLE] you cannot get your own [INAUDIBLE] unless you do them [? frame by frame. ?] I won't go into detail.

AUDIENCE: Ramesh, do you have to apply this to each channel, the [INAUDIBLE]?

RAMESH RASKAR: Yeah. You have to do it for every color channel. That's right. And from there, you have captured it.

So let's look at the [INAUDIBLE] of how exactly this works. And there's a [INAUDIBLE] explanation that many of you have seen. And I'm going to give you a very simple, intuitive reasoning of [INAUDIBLE].

So let's come to this problem of replacing a pinhole with this new type of mask, which we call heterodyned mask. So here we have a pinhole and now this part again, a pinhole. And we have nine values coming in. [INAUDIBLE] minus 4. [INAUDIBLE] plus 4, 4, [INAUDIBLE].

And then we do the next [INAUDIBLE] over. And this one's doing the same thing again. And by using this trick, you're not going to get more resolution. You're still have some-- miss some information in the [INAUDIBLE]. It's just that we don't lose the light. We just lose those spatial frequencies.

So small digression, and we're going to talk about something called [INAUDIBLE] multiplexing. And the problem is really easy to explain in this [INAUDIBLE]. Let's say I give you-- there are several bags with different weight, weight 1 to weight 9, and I tell you to weigh them. And at the end, I want a solution of each-- what the weight of each of the bags is.

[INAUDIBLE] the weighing scale, put each of the bags one at a time, and you have a solution. What are some other ways you can put it? If you don't want to do one at a time, what can you do? Sorry?

AUDIENCE: Do all of them.

RAMESH RASKAR: Do all of them? But that doesn't give you [INAUDIBLE] composing of nine. You must make nine measurements. I'll give you that. You must make nine measurements.

One choice is, you just put one bag in. And you have nine [INAUDIBLE]. But let's say your scale actually-- so let's say the weight is from 0 to 100. I'm sorry, the weight itself is 0 to 10.

But let's say your scale doesn't work very well in the first couple of [INAUDIBLE] statement. So the scale works pretty well when it's in the radiant [INAUDIBLE]. But it doesn't work so well when the weight is too low or weight is too high. So what you want to do is stay in the [INAUDIBLE].

And cameras work the same way. That's why I'm not doing this one. If the light is too low, if the light's too high, the camera cannot figure out where exactly it is. But if the light is between-- if the light is going from 0 to 55, ther when [INAUDIBLE] to 200, the camera works nice in a straight line, straight line [INAUDIBLE]. When it's too dark or too bright, a camera cannot handle it as well. And this similar analogy [INAUDIBLE].

So one solution is to do one at a time. The other solution is I can put a group of them. Let's start with [INAUDIBLE] number of three. Let's say I have three of them. I could put one at a time or I could put two at a time. Plus, I can do w1 plus w2, then I'll do w2 plus w3. And I would do a w1 plus w2.

In fact, there's three measurements. From the three measurements, I can figure out what each of these [INAUDIBLE] are. And you can actually write this down as a [INAUDIBLE] system. I will say that [INAUDIBLE] as well.

So let's say when I put the first two, my measurement is on 1, 2, and this one, 3. And all we have done here is we have said, I'm making measurements-- and 1, and 2, and 3. What I would like to know, actually, is the weights-- w1, w2, w3. But the way I have measured it is the first one is w1 plus w2 and [INAUDIBLE] w3.

Second one is w2 and w3 [INAUDIBLE] w1. And [INAUDIBLE] that's all [INAUDIBLE]. And we're just following the system. And it can tell you, from looking [INAUDIBLE]. And especially for this, [INAUDIBLE] purely inevitable.

So for three, this seems like a very good solution, because now we are staying somewhere in the middle range. We are not at the bottom range of this weight. And if you put all three of them together at the same time, you might go very far over here, as well. [INAUDIBLE] because we did the sum of two [INAUDIBLE]. So this is very convenient.

Another part of multiplexing is basically the same concept. Instead of three [INAUDIBLE] 9 or any such measurements, I'm going to basically take about half of them randomly. And I take about half of the bags, put them together, take a measurement. Then I will put them apart, again take some other half, and do our fun measurements, and so on.

See, if I have nine of these, I will create a matrix that looks something like this. I will have-- I will measure my grades that we want to go [INAUDIBLE]. My measurements are [INAUDIBLE] 1 through the 9. And what will we have here? What's the [INAUDIBLE]?

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: Sorry?

AUDIENCE: Nine [INAUDIBLE].

RAMESH RASKAR: Nine [INAUDIBLE]. And this way, we put down the numbers here, we've got to put some random numbers here. Instead of putting all the bags that [INAUDIBLE] all but [INAUDIBLE] part of this. Let's head back. If I just put one bag at a time, what will this mean [INAUDIBLE]?

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: It will just 1, 1, 1, And also the number 0. That's [INAUDIBLE]. And same thing will be going along the diagonal. [INAUDIBLE] out of the nine, I will have at least four or five [INAUDIBLE] placed within-- put randomly 101, 10-- 4, 5, 6, 7, 8, 9.

So I put-- I don't know. 01, 02. That's one sequence. And then I'll take the next one-- 01, whatever's next. I have nine such [INAUDIBLE] sequences with either four or five of one [INAUDIBLE]. And this is how [INAUDIBLE] sequence as [INAUDIBLE].

So it's sort of taking one measurement at a time. But you will take a linear combination, in this case just a sum of our [INAUDIBLE]. And the benefit of that is that you're staying in a range that's [INAUDIBLE].

I used the analogy of this one with where we're going. Using a pinhole is like doing one measurement at a time. You're only measuring one ray at a time. Using a lens [INAUDIBLE] is like putting all the bags at the same time.

And what we will do now is place a mask where only about half the values are coming through. So how many variables? So you're gonna be putting on half the input variables, and measure those. [INAUDIBLE] analogy, where it is going, and I'm able to predict what the variables will be.

So let's focus on this one slab here. If you just focus on one slab here, light comes in 2 and 3. Instead of 9, we'll say, 3

This one's just blocking this light here and here. What I will do is I'll block it from certain directions. So let's say I will make this-- actually, 14 doesn't look that interesting. So let's go to [INAUDIBLE]. I'll make it 1-1-1-0-1-0-0. [INAUDIBLE] in our half.

What's going to happen is, for a certain part of the image-- if you look at the pixel here, I shoot the-- this is my lens. This is my pixel. If I shoot the ray here, it's blocked. But if I shoot the ray here, it goes through. If I shoot the ray here, it's again blocked.

So what I'm getting at this pixel is not sum of all the rays or some image of [? conventionalist. ?] If I go to the next one, I get some other combination, because for this one this one was blocked, but for this one this goes through. Same here. But now this one's blocked, and so on. Because of the displacement, the combination that we're getting is evolutionary.

And once we have done the [INAUDIBLE] combination, which is what we have here, about half the light goes through. But what we have seen here is the linear combination. We have these measurements. We don't have the original radiuses. Just simply calculating it will recover these intensities. And there you go. We're back to a traditional [INAUDIBLE].

So that's a very easy way of-- very intuitive way of thinking about how we can use [INAUDIBLE] multiplexing with [INAUDIBLE]. Again, the solution. So we get now half the light. [INAUDIBLE] but this lasts half in here, half here. In this case, we just have on half of the [INAUDIBLE] and [INAUDIBLE]. So we are [INAUDIBLE] for the masking and the actual camera is ready to go.

Fortunately, you have to do a lot of computation on this equation for the whole image. Image is 16 megapixel. This matrix could be 16 million times 16 million if you do it in a brute force fashion. But of course, you just [INAUDIBLE] simplest conversion [INAUDIBLE] based on this that make it really, really important. So we won't go into inversion, but we want to stay in problem as [INAUDIBLE] would be and say that it's possible [INAUDIBLE]. Yes?

AUDIENCE: Can you process other one or is something [INAUDIBLE]?

RAMESH RASKAR: That's a very good question. So what I show you, I kind of cheated. I showed you a mask that looks like cosine. And what we're doing, it was certainly realized just in the last one year, is that [INAUDIBLE] and cosine is one and the same thing.

So if you take a bunch of cosines-- and [INAUDIBLE] information, by the way. So that's why we have it here. So if you take cosines of different frequency-- so this is like taking [INAUDIBLE]. And you take a [INAUDIBLE] take positions on the frequencies. What we're doing is we're projecting on, again, different [? carriers ?] and so on.

So if we do [INAUDIBLE]. If I do cosines, and then I'm going to make them sharper, sharper-- if I place all of them together, you start getting that [INAUDIBLE] in the center. And you get a bright spot. And away from it, you get a dark spot.

The sum of all of this, as you can imagine, ends up being something like [INAUDIBLE] and then [INAUDIBLE] and then repeat some variation. So that's the part [INAUDIBLE]. And it turns out we can place these cosines in such a way so that we actually get a [INAUDIBLE].

And [INAUDIBLE] really takes a lot of work. We'll talk about it, about how you can actually get binary [INAUDIBLE]. And so something that [INAUDIBLE]. Because printing a binary mask has been measured. It's more convenient than printing a mask where [INAUDIBLE] is changing [INAUDIBLE].

AUDIENCE: And when you phrase it as a linear system, there is no-- I mean, no different, no sense in having-- blocking half the light or so. I mean, partially, with a partial transparency, but you could just go binary anyway.

RAMESH RASKAR: Yeah, exactly. If you go binary--

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: You're still going to lose half the light.

AUDIENCE: Yeah, yeah.

RAMESH RASKAR: [INAUDIBLE] some of this.

AUDIENCE: Yeah. But I'm saying, instead of using [INAUDIBLE] and say 0.75 or so, I mean, [INAUDIBLE]

RAMESH RASKAR: That's a great point. So let me rephrase what this fellow over here is saying. He's saying instead of ones and zeros, which means the sum of all this is still half of the total light, why not make this 1 and 0.75, and 1, 1, 1.75? So you still have this variation, but most of the light is going through.

And it's actually a great topic for research as to what exactly this pattern should be. And for us, binary was very convenient on 0. And as you play with that parameter space, you get many, many different interesting solutions. Some of them are geared more towards getting to the photo, and some of them are geared more towards capturing the lightfield.

So for example, if this is all ones, then you [INAUDIBLE] your photo. If this was all ones, that means every point is transferred, which means there's not much [INAUDIBLE]. If you put all zeros, that means it's [INAUDIBLE] go through. If you use a [? spectrum, ?] then it's a random [INAUDIBLE]. But if you use any [INAUDIBLE] in between, then [INAUDIBLE] optimizing [INAUDIBLE]. Yes?

AUDIENCE: I guess one [INAUDIBLE] equations [INAUDIBLE]. Because holding [INAUDIBLE]. They say others [INAUDIBLE]. You [INAUDIBLE] in the Fourier [? trans-- ?] in Fourier [INAUDIBLE]. And you told me it's very easy. And it's very easy to figure out what kind of [INAUDIBLE].

RAMESH RASKAR: Right. Exactly.

AUDIENCE: And I wonder what [INAUDIBLE] similar effect [INAUDIBLE].

RAMESH RASKAR: So I really would like to do [INAUDIBLE] optical domain, but we don't have the pleasure. The light is [INAUDIBLE]. through that light. And we can only do blocking and unblocking.

So if you had a way to sense the image first, and mostly assess the image, and change your mask so that you can choose the right parameters for the scene, then what you're saying is true. So you get a softer compression. You get to look at the signal before you compress.

AUDIENCE: But I thought you could [INAUDIBLE] it's not focused on the image. The image won't have any high level [INAUDIBLE]. But [INAUDIBLE] and you cannot capture frequencies like that.

RAMESH RASKAR: So frequently.

AUDIENCE: Exactly. And it doesn't [INAUDIBLE]. But [INAUDIBLE] we can [INAUDIBLE].

RAMESH RASKAR: Exactly. So what [? Deena ?] is just asking is, if I know that a machine is [INAUDIBLE] 200 pixels and 9 variations-- so let's say for the photograph, what I'm going to do is put a photograph here and then capture its lightfield. And I'm going to [INAUDIBLE] 200 pixels. And I have nine views for each of those pixels. Then what I've done is built the correct system.

But let's say somebody gave me a photograph which had, actually, 400 pixels. And, let's say [INAUDIBLE] 300 pixels, and then only six views. I don't know the priority. But you just note there is a 200-pixel photograph or a 300-pixel photograph.

I will continue to decode that as a 200-pixel photo. But the problem will be, because this team has higher frequencies, different areas, it's the same question you're asking. And you have no idea of knowing [INAUDIBLE].

So a typical solution in any signal processing, any device that has been sampling, is you prefigure. You reduce [INAUDIBLE] by [INAUDIBLE] or cutting off high frequencies so that it will not [INAUDIBLE] as low frequencies.

In the optical domain, unfortunately, it's very difficult to do. [INAUDIBLE] that will convert something that's 300 pixels to 200 pixels purely optics. You can do the software. You can [INAUDIBLE] or you can smooth [INAUDIBLE] software. But it's not that easy to do in [INAUDIBLE]. At least we don't have a solution. Maybe you will come up with a solution. And once you do that, then you can [INAUDIBLE].

AUDIENCE: So it says here that if [INAUDIBLE]. And so I'm thinking just about will happen to-- [INAUDIBLE] and the-- it's a complex number. And you have [INAUDIBLE] and the magnitude. Are you capturing both, or are you only capturing magnitude?

RAMESH RASKAR: Only magnitude. So the other way to think about that is, as I was drawing these cosines, instead of one, zero, and [INAUDIBLE], what you're doing is really truly a prediction. So the very first frequency we have is first cosine.

The second one we have is the highest [INAUDIBLE]. The third one is that you will have the lowest cosine. And we have these nine cosines with four with plus, negative phase, and other four with [INAUDIBLE]. And the middle one is just [INAUDIBLE]. That's why I [INAUDIBLE] this particular pattern. And it turns out, instead of using cosines, we can do the binary [INAUDIBLE].

AUDIENCE: [INAUDIBLE]

RAMESH RASKAR: That's the keyword [INAUDIBLE].

AUDIENCE: [INAUDIBLE] because now we [INAUDIBLE]. And are all these Fourier transformed different? Or are some of them [INAUDIBLE]?

RAMESH RASKAR: No, no. They're very completely different. They're going to be [INAUDIBLE].

AUDIENCE: So [INAUDIBLE]?

RAMESH RASKAR: Because, remember, I'm taking cosine and it's [INAUDIBLE]. So basically, again, I have nine unknowns and nine measurements. And I can't come up with any of these signatures as [INAUDIBLE]. I can mix them, bunch of ones and zeros. Just [INAUDIBLE] missing. Or I can do [INAUDIBLE] some kind of [INAUDIBLE] prediction. I can do [INAUDIBLE].

AUDIENCE: From the way you described putting the mask there, the way this matrix is formed, it seems like the second row is probably a shift of the first row.

RAMESH RASKAR: It could be, because [INAUDIBLE]. So that's the [INAUDIBLE].

AUDIENCE: Because invariably, inversion probably can be easier, because you know the pattern of the matrix. And it's not a brute force inversion. Maybe you can do [? automation. ?]

RAMESH RASKAR: Exactly. So all the beauty and elegance of the effort comes in in choosing the right type of mask and using-- codesigning a decoding scheme that's-- it doesn't amplify the noise, and so on. You had a question?

AUDIENCE: I was going to ask [INAUDIBLE].

RAMESH RASKAR: That's the same question that [? Deena ?] is asking. In fact, I'll-- anything beyond 200, so the-- from 201 to 300 doesn't dominate the picture. I like to think of it as script domain because [INAUDIBLE] and because [INAUDIBLE].

AUDIENCE: [INAUDIBLE] from a little while back. Can't we use something with [INAUDIBLE] arrays [INAUDIBLE]?

RAMESH RASKAR: So it turns out [INAUDIBLE] whether you think about them as 4D real space or you think of light as wavelengths [INAUDIBLE] for on measuring the [INAUDIBLE].

AUDIENCE: Yeah, I was just thinking I could guide [INAUDIBLE].

RAMESH RASKAR: [INAUDIBLE] down from UCLA?

AUDIENCE: Yes. [INAUDIBLE].

RAMESH RASKAR: Right. But he's using diffraction for that. And we'll come back and talk about this. And we'll talk about [INAUDIBLE]. So a lot of you are wondering, where does this all fit in? This seems like a lot of math, a lot of algebra, or something you could just put a pinhole in and be done with it [INAUDIBLE]. So for example, right here, I am going to explain to you how the same exact technique can be manipulated in very different ways in two very [INAUDIBLE] series [INAUDIBLE].

AUDIENCE: So just a quick question for that. So have you already compared the results using this cosine mask with the binary ones?

RAMESH RASKAR: Yes. Yes. In fact, here, let me show you.

AUDIENCE: Because in theory, they should be equivalent. But in practice, because maybe some high frequencies that you introduce by the binary mask coming in-- and then you have to truncate at some point.

RAMESH RASKAR: In simulation, they're identical.

AUDIENCE: Yeah.

RAMESH RASKAR: But you're right, because what isn't happening is when you print these cosine masks, your intensities are actually not transposed [INAUDIBLE] very well in cosine. But it's actually wearing something like a step function and so on. Because this digital levels by completing this [INAUDIBLE]. So that ended up creating a lot of trouble.

AUDIENCE: And for the binary one, as well, because then you're going to have some sharp transitions that, in the Fourier domain, they may go all the way to high frequency.

RAMESH RASKAR: Exactly. It's a conditioning of the [INAUDIBLE] transitions not right. And your model is incorrect, [INAUDIBLE]. So the calibration is very key.

But the one benefit is that because now we're using masks rather than [INAUDIBLE], you don't have to have the lens [INAUDIBLE] so that image of the lens is created as a sharply focused version on the sensor. In other words, very, very, strong constraint that we must put distance [INAUDIBLE] so that the image of the lens is formed on the sensor. We don't have that anymore.

Even if we have some [INAUDIBLE], there's some [INAUDIBLE], some tolerance in [INAUDIBLE] of the mask. Because all we're doing is we're taking the values we want to measure. We're going to mix them around and get new measurements and [INAUDIBLE]. That gives us our freedom.

So if you go back to the analogy of nine bags, it's rough-- it's roughly a very specific combination of those nine bags. It will be a slightly different combination of the nine bags. My mixing matrix [INAUDIBLE] most efficiently, but it would be good enough to do a reasonable job of [INAUDIBLE]. So that gives you a lot of flexibility [INAUDIBLE]. So now I'm going to show you how we can use this [INAUDIBLE] technique.

AUDIENCE: So, Ramesh, but it seems that the binary mask made to do some diffraction artifact, right?

RAMESH RASKAR: They do. They do. Are you going to talk about that?

GUEST SPEAKER: I'm not sure I got to that. That's fine. You can see the results, and then we can talk about [INAUDIBLE].

RAMESH RASKAR: Yeah. And, just-- you can just go slow, because we just talked about it [INAUDIBLE].

GUEST SPEAKER: OK. So I--