Image Single Processing

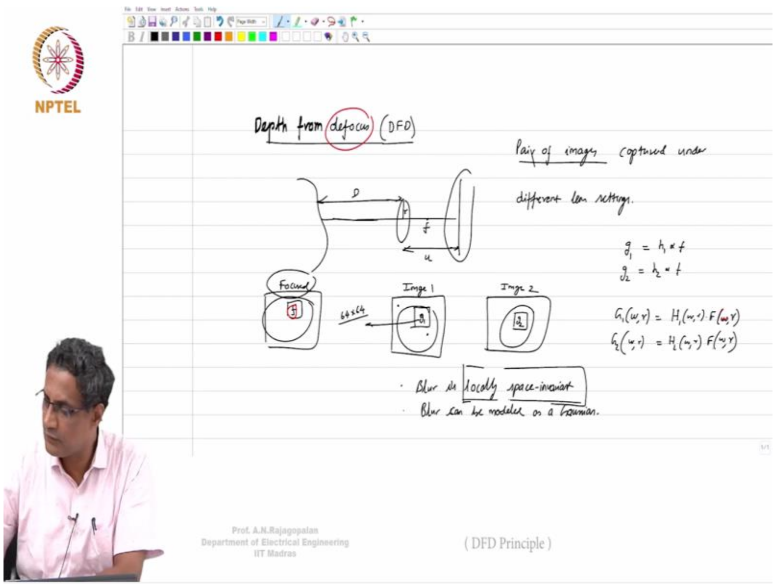
Professor. A. N. Rajagopalan
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture No. 34

DFD Principle

So, for today will be the last class on this topic and tomorrow we will move to, move to another module, which is an Image Transforms,

(Refer Slide Time: 00:28)



So, this topic is actually Depth From Defocus this is too wide. It is called a DFD as suppose to, suppose to shape from focus, this is some more like Stereo, but then except that in the sense that you get you have to acquire a pair of images just like you do in Stereo but in the main sort of a difference is that you do not move the camera here there is a Panasonic Lumix some GH5 or GH6 that came out last year I think 2018 I think so now that camera uses this.

In order to do the hunting so I thought we might as well look at it so that you will at least know something about it because they do they do mention that the 100 times faster because of this. So, I thought it is it is worthwhile spending a little bit of time. It would not take but then for some approximate calculations like for hunting and all that it is quite good and so you will see why?

So, the way this works is when you have read you have this kind of a 3D scene and then you have a lens, of course this all, this all say a real aperture camera. And do you have an image plane? So, what you do is you take one image, so, so actually this, so you need a pair, it captures a pair of images, you can cap you can also capture more than, more than 2, but minimum 2.

Like I said, if you just capture one then I think it is too hard. Of course even though now deep learning techniques and all there that do with just one image and so on, but those are all training based in. So, we do not want one of those kind of things now, so pair of images and so you also captured, such captured and the under, captured with rather different lens settings.

Lens settings, what that means is? You might actually know you might, you might change the aperture of the lens, whatever let us call this R, you might change the aperture so that means you will capture one image with some with some lens setting. And then you capture a second image by changing the lens setting or you could vary one or more of these more of these parameters.

So, one is R another is another is this image plane to lens distance, but you do not the lens you just move the image plane. And I am going to leave it to you to show that if you move the image plane there is no parallax, camera center does not move, you can move the image plane even is the scene is 3D. It is not, it is not a parallax. I leave it to you. Then, then what else can you and then the focal length?

So, so you can change one or more of these lens settings, it assumes that that it knows them exactly what is this Lens parameters are, kind of assumes that all those are known. That is why I said it makes some assumptions that if you can, if you made those assumptions are are valid in the sense that if you can satisfy those assumptions, then yes, then you can use this. But again, there are some other some other things that I will tell you.

So, what it does? It is again, it is again, a very, very fast technique. It is again, a parallel technique just like your shape from focus, except that it can just work with a pair of images. One of these one, so how would so imagine that idea captured Image number 1 here, and then you are captured another image, which is Image number 2. So what has happened is because, because the scene has not moved at all, and we have not, we have not moved the camera, moved the lens.

We have not, we have not moved the, moved the lens or anything if at all, we can change U F for R. So therefore, what it means is that, that you get, let us say some space with a blurred image depending upon what the 3D scene is like. And then again, in the second one, you will get again,

again, another kind of space, variantly blurred image, depending upon what you do. For example, if your decrease R, then it means that your blur will uniformly go down all over the image, and if whatever.

So, so there is a function of what you do, do with these three things, you can just choose to vary just 1 or 2 or maybe all 3. Now, the point is this. So now, what it what, what it what it does is it tries to compare these two images in order to be able to tell what is the depth because the depth is still the same, because from, from here to here, at some, some point from the center is at is at a distance D, then it remains at D because you are not you are not moving the lens at all.

So, then the second image also remains at D. So basically D is unchanged in both the images and if you move the image plane you can still do do do a compensation for that because image plane if you move this guy in the sense that if you move this to the left or to the right you can do that and then still have a compensation the sense that you can kind of align the two. See the point is now, the now what, what is going to be done is the following.

So, the assumption that it makes is that blur is locally, locally space invariant. Blur is locally locally space invariant. So, this is like saying that, overall, the space varying so it is spatially varying. So, if you, if you if you kind a compute what is the kind of blur that you have here, maybe that is not the same as what you have here that may not be the same as what you have here.

But over a local region, suppose a pick some pick some kind of a patch, but these patches are typically small. They are not typically more than 64 by 64. So, over a 64 cross 64 patch for example, I am saying this as an example, take a 64 cross 64 patch, then you would actually assume that within this patch, the blur is kind of a constant, which may not be true because it could be slowly varying.

It may not actually be true, but that is an assumption that it makes, because it makes an assumption just to just to simplify matters. And you will see why? Another assumption that it makes us the blur can be modeled as, as a Gaussian, blur can be model, which is also the reason why this is why you know why it says from defocus. The other one was from shape from focus on shape from focus we never, we never, even modeled the blur on it.

We simply said when does a point come in focus, we never worried about what might have been the shape of the PSF and all and the rest of the frames we never, but here, it does, it does model blur. So, it says blur can be model, model the approximately that will that is what it means. We can approximately model this as a Gaussian, which again means that you could, you could actually incur, incur some kind of error because of these approximations that you are making.

But no, it turns out that it turns out that in spite of these approximations, this is still for certain applications, it may still be, it may still be quite useful. So, the way it works is as follows. So, if we are going to look at a patch here, let us call this G1 and let us kind of look at the corresponding patch here which is G2. Now, now, now, what you do is So, if you look at if you look at G1, so, so, you know that there must have been some underlying focused patch there, which got blurred because of the because of the lens parameters.

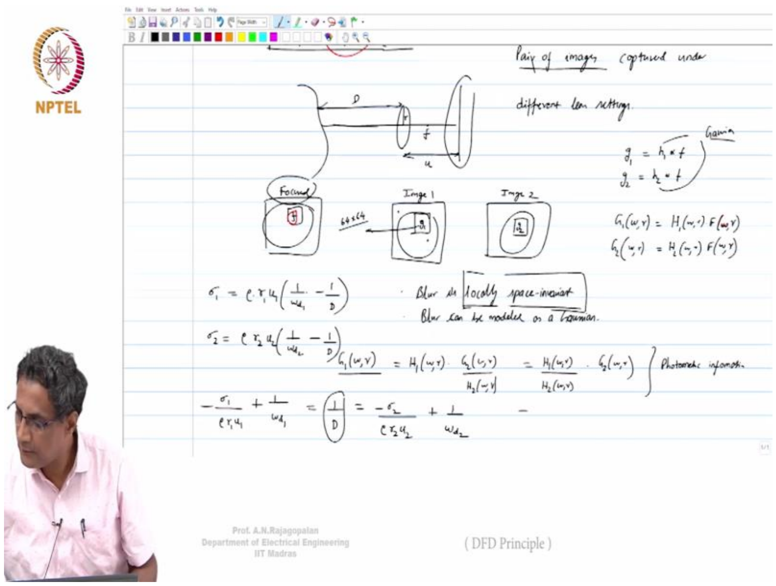
I mean, ideally, there must have been something that was all focused which should be, which should be a pinhole can have an equal and we can think of some focus here, which we cannot see of course, but this focus guy is probably what is originally there, but then we cannot see that at all because our lens is not able to bring this 3D scene into focus. And that is what we are in fact using it using as a Q2 no find the depth.

So, assume that assume that there was all over this image, if this guy was actually focused all over then under this patch, that you can assume that there was some F, which got blurred by some, let us say a PSF H1 to give you G1, then when you change the lens parameters, it gave you something like an H2, which got convolve with F in order to give 0 because convolve because of this local space invariant assumption.

Because of that, because of that assumption that we are making upfront. So, what this means is that you can write this as H1 convolved with F1. And then you can write this G2 is, this is a very simple, simple method G2 convolve with F2. So, that should not take too long to do this. So then if you kind of go to the, if you compute DFTR or something, then, then you can look at something like G1 enough, omega comma some, some Nu, then let us say that is like H1 omega Nu into F1 of omega Nu.

Initially, we will take a continuous kind of case okay and then a discrete approximation is what you will typically implement. So, G2 omega Nu then H2 of omega Nu and into F2 of omega Nu both is F this is not, not F1 F2 because underlying patch is the same no cannot write F1 F2. See this guy is F. See this is F and F blurred with H1 is G1 yeah so, this is not, this is not F1 and F2 those H1 convolved with F is to convolved with F. So, this is like F of omega, like F of omega Nu this F of omega Nu again.

(Refer Slide Time: 09:44)



So, then it follows that that so you can kind of light again. So, G1 by G2 let us say for or, or we can write G1 omega Nu is equal to H1 of omega Nu and then substitute for substitute for F of omega Nu which is like G2 of omega Nu divided by H2 of omega Nu. Now, now whenever you do a division it will let us assume that that this guy is non 0.

Because if it is 0 for some omega Nu then you cannot strictly speaking divide okay which is why I said that a continuous case that we will assume that it is, it is a Gaussian therefore, if it is a Gaussian then, then it is to kind a divide, goes to 0 only, only as omega tends to infinity. So, so it means that even omega is not H1 so F is simply G2 by H2 or this like H1 by H2, H1 omega Nu by H2 omega Nu okay into G2 omega Nu.

This are the so, we have something like this on the one hand, which comes from the intensity information that is all the photometric information so this, this is coming from the, from the photometric information. We can we can simplify this further but we will come back, come back to that in a minute. There is one more thing which is happening, which is like saying that no since we are modeling H is H1 and H2 these are all being modeled as a Gaussian.

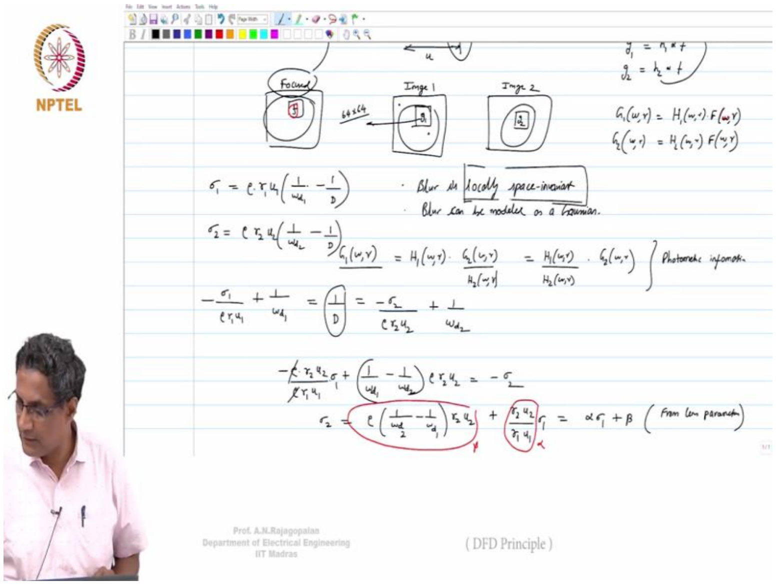
So, in each case where you have something like a sigma 1 parameter and then for the other, other you have some like a sigma 2 parameter. So, we know that sigma 1 might in some row. So, let me write this is r1 u1 1 by omega D1 minus 1 by D. So, that basically omega D1 will again be something which gets automatically governed depending upon how you choose your choose your focal length r and v, v or u and then sigma 2 that means when you change the lens parameters row is a calibration constant is the same camera row cannot change.

So, you have row into r2 u2 1 by omega D2 minus 1 by D. This will be, this will be your, this would be the second equation. So, one for the first sigma another for the second sigma this is all local. So, the sigma 1 and sigma 2 could be varying all over the image. So, you are sitting at one patch, and you are sort of saying that this is the sigma that governs that patch.

Then, then what we can do is then we can write sigma 1 by a row r1 u1 is equal to 1 by or we can say minus of this plus of 1 by omega D1 is equal to 1 by D. And that is also equal to what is this, row r2 u2 by omega D2, wait a minute let us pull this here so, minus of sigma 2 do the same thing row r2 u2 minus of this plus 1 by omega D2. And the main assumption that you are making is that that D is the same which is valid in this case because you have not moved the camera

So you say, you say D for that point only, only the only the amount of blur that that point suffers on the image is changing, but the point itself is still the same on the scene. So, which then means that, you can you can maybe read multiply. But let us do some row r2 u2 to do it in whichever way you want.

(Refer Slide Time: 13:22)



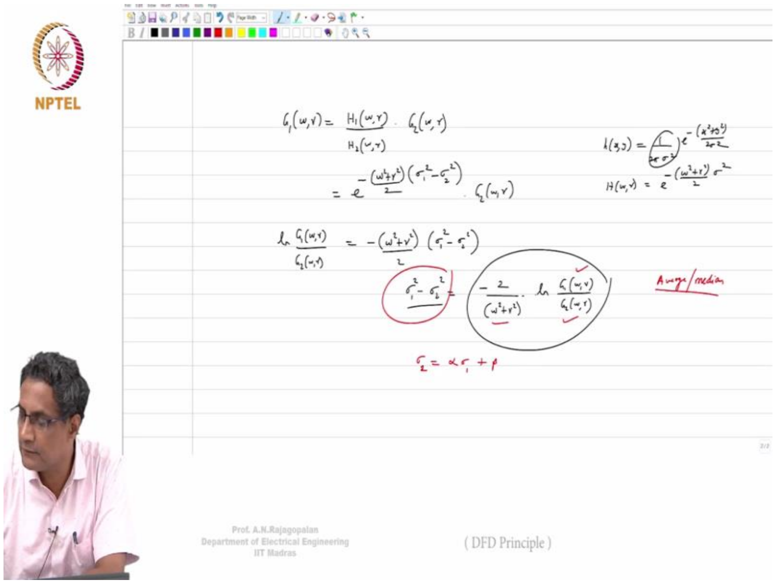
So, we can say, from here, we can say minus row, row r2 u2 by row r1 u1 plus 1 by or plus 1 by omega D1 minus 1 by omega D2 into row r2 u2, r2 u2 is equal to minus sigma 2. Or we can say sigma 2 is equal to 0 will cancel off or we can say sigma is equal to row into 1 by omega D1 minus 1 by omega D2 into r2 u2 plus omega D2 minus D1 here.

So, I should write it as D2 minus omega D1 because I want to write everything as positive, then this is plus r2 u2 by r1 u1. So, we can write this as alpha sigma 1, so, suppose we call this is alpha, let us call this as alpha. So, alpha sigma 1 and then plus let us say beta, where we can say we can we let us call this guy as beta.

So, this is beta, let me check this if everything is okay. So, sigma 2 is 1 by omega D2 r2 u2 plus r2 u2 by r1 u1 sigma 1. So, what this means is that, that that you have this purely from, from the, from the left from the lens parameters. So, this is purely coming from coming from lens parameters because you change the lens parameters, you have one equation that actually that that kind of say relates sigma 2 and sigma 1.

Now, let us, let us get up go to this one, G1 is equal to H1 by H2 times G2. Let us go to the photometric thing which is the intensity.

(Refer Slide Time: 15:45)



So, what was it? So, we had like G1 of omega Nu is equal to H1 of omega Nu by H2 of omega Nu times G2 omega of Nu. So, then, then we know that H of omega Nu is e power minus omega square plus Nu square by 2, but in sigma square gets multiplied on top when you take the Fourier domain. Then if you have the H of xy to be 1 by 2 pi sigma square e raised to minus x square plus y square by 2 sigma square then you take the Fourier transform and what is that as long as you have this factor in it becomes e raised to minus off omega square plus nu square by 2 sigma square.

So, the sigma square comes on top when you go to the go to the Fourier domain. This is something that you all know. So, the so, this Gaussian is one of the special cases where they the shape is retain when you go to the Fourier domain the shape remains the same still remains a Gaussian. So, then this means that this leads to e power minus of omega square plus Nu square then we can write this as sigma 1 square minus sigma 2 square into G2 of omega Nu.

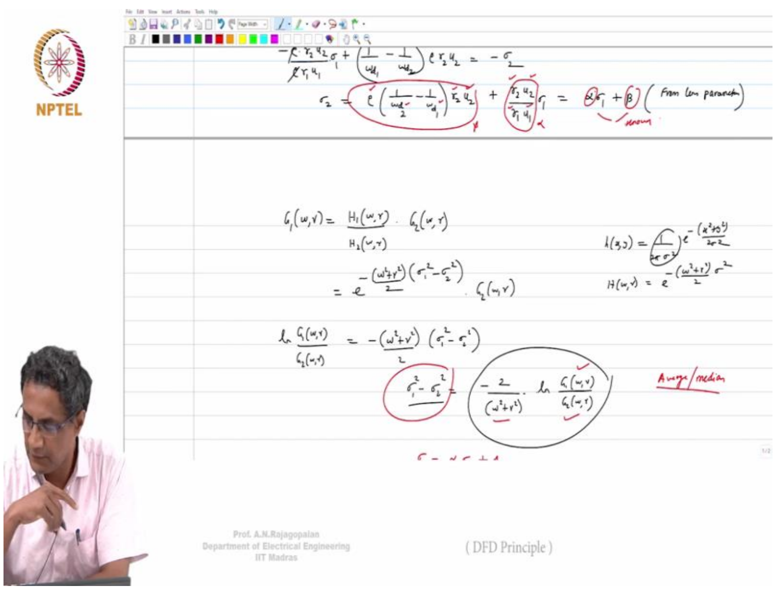
Or we can say ln of G1 of omega Nu by G2 of omega Nu, I will bring this down is equal to minus of omega square plus Nu square by, there should be by 2. So, Nu square by 2 into sigma 1 square minus sigma 2 square. Or in other words you have sigma 1 squared minus sigma 2 square is equal to minus of 2 by omega square plus Nu square to ln of, G1 omega Nu by, so you get an equation of the type.

So, so now the point is this so actually, this is supposed to be valid for all omega Nu. Of course, we may not take omega Nu and both 0 because then it will blow up but otherwise, for all frequencies, it looks like there is only there is only one value that you should ideally get, whatever I do for any frequency, it looks like theoretically I should get only one value, but in practice, what is some there is something called leakage and all that I do not want to go to that.

So, what is normally done is you take the, we take the average of this of this value on the right hand side, because this is this is available to you know, so, you know, you know the value, you know, you know, G1, you know G2, you know, at what value of omega Nu you are computing it, so the entire thing on the right hand side is known to so you can either compute the, the average or you can compute the median.

Again, it is up to you average or median, you can compute in order to get an, get an estimate for sigma 1 square minus 1 because actually, this is a very simple thing. It should work for all omega Nu. But then you do noise and other things. And there is something called leakage and all that. So, you may just want to average. Now, you have actually the other equation where you had sigma 1 is equal to alpha sigma 2 is equal to alpha sigma 1 plus beta that is what you had no alpha sigma 1 plus beta where again.

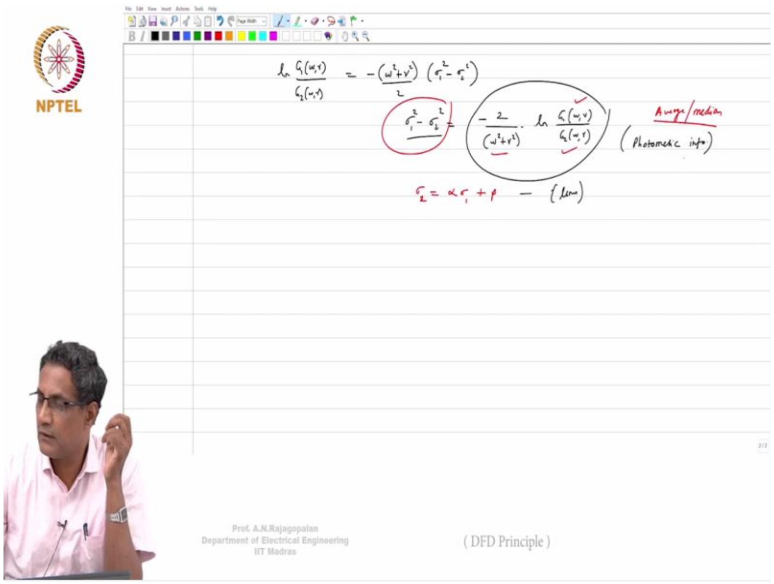
(Refer Slide Time: 19:16)



Here if you notice, these things are all known okay, we are assuming that the camera calibration So, if we look at beta, if we assume that role is known omega I mean this is like only these things are known r2 u2 all these, are all these are related to lens similarly r2 u2 r1 u1 these are all related to lens. So, we are saying that alpha and beta this should be known.

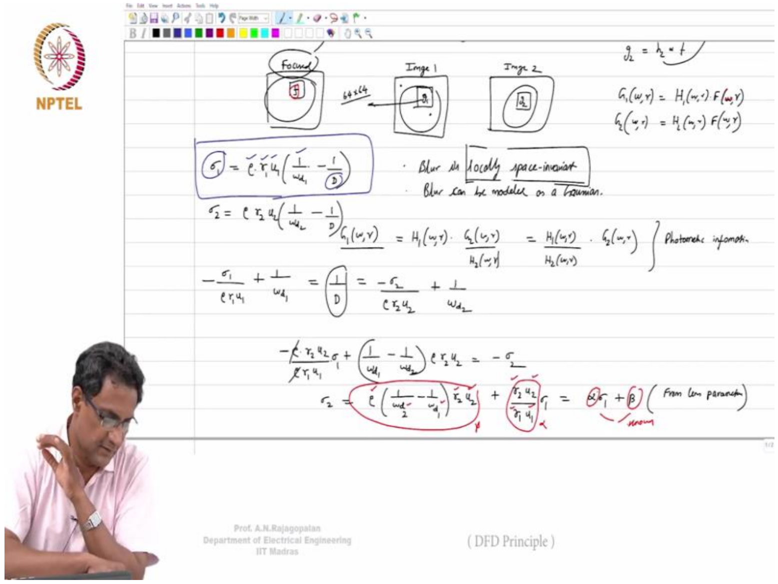
Again, if you do not know then of course, you cannot, you cannot use the method. So, so, you are assuming that all these things are accessible, and you actually have this I mean you can, you can gain access to these values.

(Refer Slide Time: 19:50)



In which case, so, in which case, you will actually end up end up end up with one equation which is which is coming from the lens. The other equation which is coming from the from the photometric from the photometric info and then and then and then you can actually solve for it, it is all for sigma1 and sigma 1 or sigma 2 or both and then once you know sigma 1 or sigma 2 then you can actually go back and then if you know everything else, you can actually go to this equation.

(Refer Slide Time: 20:22)



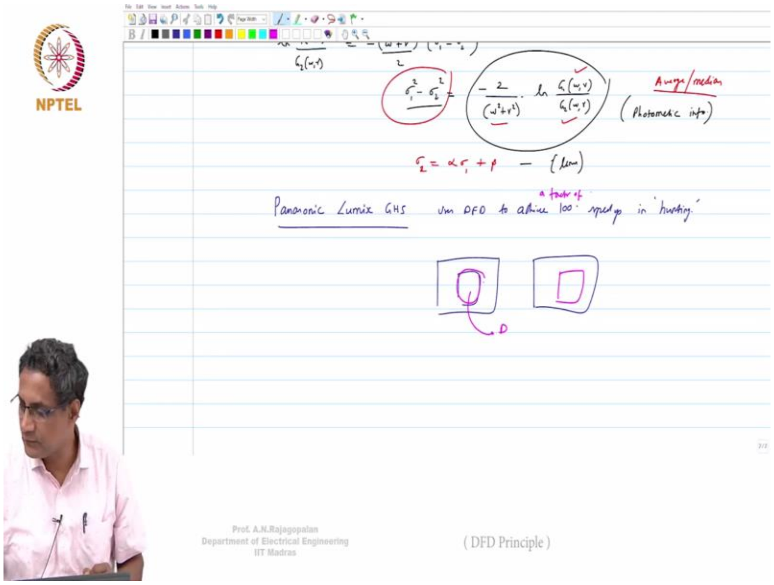
The first equation that you had that was relating depth and so you had this equation, so you can go to this equation, plug in the estimated sigma 1 or sigma 2, and then know all the lens parameters, plug them all in, and then, then that would give you a given estimate of depth at that location, I mean, that will be the depth for the center of that patch by the way.

That is be the debt for the center of the patch, then you have to get a move this patch to the next pixel and then and then you can just go on all over the image and the but again, this can also be completely parallel because you are not going to you are not putting any constraint on a adjacent depth value should come out to be similar or anything.

If you do not put anything like that then it can be very fast but it can also be noisy. Now these guys this the Panasonic guys what they have done is they have done something smart so they know that the (21:10) is probably not the not, not the probably ideal thing to not get a depth map and they do not know need a depth map.

Depth map by which we mean the entire scene we want to know where is every point that is not what is of interest to them. So, what they have done is they have used it to naturally do hunting.

(Refer Slide Time: 21:25)



And so this hunting thing so this is Panasonic you can check this for those now you should understand if you go and look it up Panasonic Lumix I think it is called GH5 or 6 I do not know okay, something so this uses DFT to achieve 100 percent speed up over let us say existing speed up in hunting. So, the hunting I told you already and so this is something that uses for auto focusing so what, what it imagines this.

So instead of capturing, like we like I said normal typical cameras will capture about under 50 to 100 frames and try to look for some contrast measure, which will tell right where which lens setting is this thing coming into focus. Instead of working on that many images, what they do is they simply capture two images by changing the lens setting, they, they assume that the center of the image is what is of probably interest to you.

Naturally, that is what I said even the last lesson, you may be interested in the center, and they assume that the, that the central thing is probably one flat plane, it could be wrong. It could be that something is changing very drastically there, but this is all assumptions that they make. So, so what they are assuming is that is that that this patch is almost has one sort of a constant depth, okay. And then, and then what they do is they will actually compute uses DFD, and then because it is only two images now, and they will they will not get an estimate of depth for this for that layer.

So, it is like a layer that I can think of the scene as being a layered scene and you are interested in the middle guy, the central portion of the image. Therefore, for that you find out a depth value. And once you know the depth value, you know that this may not be very accurate, so what they do is around that depth they capture. So, now that they know where the scene is likely to come into focus around that they just capture a few frames, 10 frames or something and do a shape from focusing.

Not shape from focus they do a contrast sort of a measure in order to then arrive at the actual incidents and this apparently it is it is very quick, gives 100 percent sorry, 100 a factor of 100 not 100 percent, it is a factor of 100 achieve a factor of 100 or 100 percent. A factor of 100, 100 percent may not mean enough, we do not care. So, speed up in hunting.

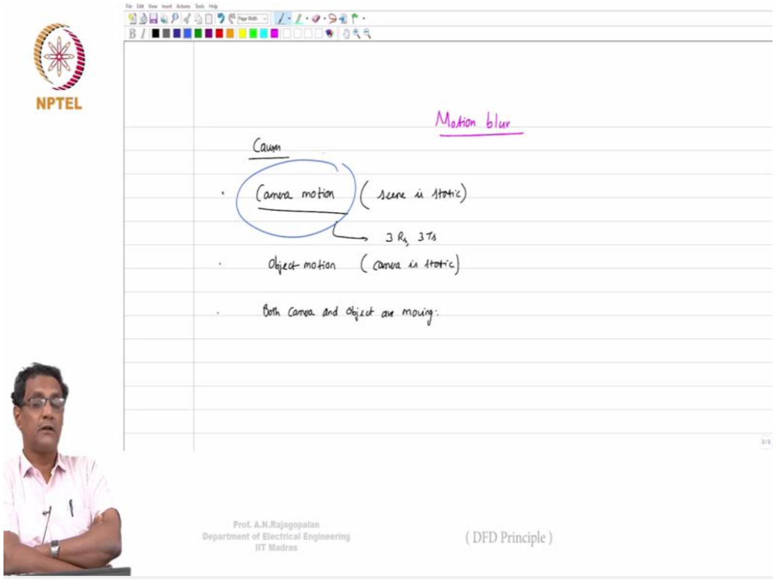
So, this I thought I will just tell you because this is like stereo and then and then afterwards, but if you see the papers that are there, they they kind of talk about stereo plus D focus on all that and then, then and then it all kind of then it becomes very, very complicated. You can have a lens and then you can move the lens then you say there was also D focusing effort to change the lens parameters.

Also translate the camera so you have a stereo D focus cube blur Q, can you can you combine both in order to come up with a kind of a better estimate? So, so those that do kind of say research in these areas will try to think, in those ways. How do you can expand the scope of this problem? Because now to, to say that, a camera has to be exactly still there and not move and then they should be able to change the lens, lens settings may not be a practical thing.

So, they would assume that what if there is a small shake or something and you capture the second one cannot assume that it will exactly be at that location, which means there will be a parallax and then, then but then equivalently, what will happen is your computation will go up will become a more complicated problem, but that is okay. If you want to write up write up paper that is what you do.

You make a simple problem, make it kind a complicated. And so you write papers. No, I mean, you do that provided it makes sense to do it. I do not mean you do not write papers just because you have to write one. Now, the last thing that I wanted to talk about was until now, we have looked at what is called the Optical Blur. Now I want to ask you something.

(Refer Slide Time: 25:01)



You have also seen, seen motion blurred images. Assume, assume that assume that the, that the camera itself, itself is not really, not really introducing any blur in the sense that the optically it is completely sharp. I have a, I have a camera which can capture all the scene completely shocked but then when I actually capture something that is moving I see some kind of a smearing effect okay that that does not happen because of optical defocus that happens because either, either the camera is moving or the what you say okay or, or you have a moving object so let me write it down.

So, this motion blur what, what could be the likely causes? So, one camera motion again no it could it could be a complete 6D motion because it was not in the sense that all 3 rotations all 3 translations, not under our control so camera, camera is camera moves, but then scene is static. This is, this, this is somewhat sort of doable the sense that you assume that the scene is static like for example this class and then I take a camera and then for some reason my hands are not stable and there is a shake.

So, you will see that all over the image there will be some blurring. And again it depending upon what kind of motion I incur my hands incur the read what you, what you will see as blur will also change across the image. So, scene is static, so camera motion can be kind of full blown in the sense that three, 3 R’s and 3 T’s. 3 translations, 3 of these rotations then the second thing could be that object motion, okay need not be just one object that could be kind of multiple objects if there are multiple objects and then it makes it even more hard.

But let us assume that there is object motion but the camera is static. And then the third thing is both a moving both camera and object are moving. Which, which, which is which is even more hard and object are moving. Now what, what I kind of intend to show is, so this is also another form of image formation. If I ask you, how do you think that that image gets formed?

Let us kind of look at the first case, which is just the camera motion, which is the simplest of all. How do you think that that just as when you have a lens, you could know you can know, they think about how the image formation probably happens? Suppose I ask you the same question now, with respect to with respect of motion blurred image, somebody gives you a motion blurred image.

How would you get to interpret this now because there is no Lens Blur now? What would be your interpretation? And then, then we will see the math and the math is pretty straightforward. But let us let us just intuitively how do you how do you interpret because the scene is not focused but because you move that is blur now. So, how would you because if you can do this, then, then it actually opens up interesting avenues. I will talk to you.