Advanced Measurement Techniques in Fluid Mechanics and Heat Transfer

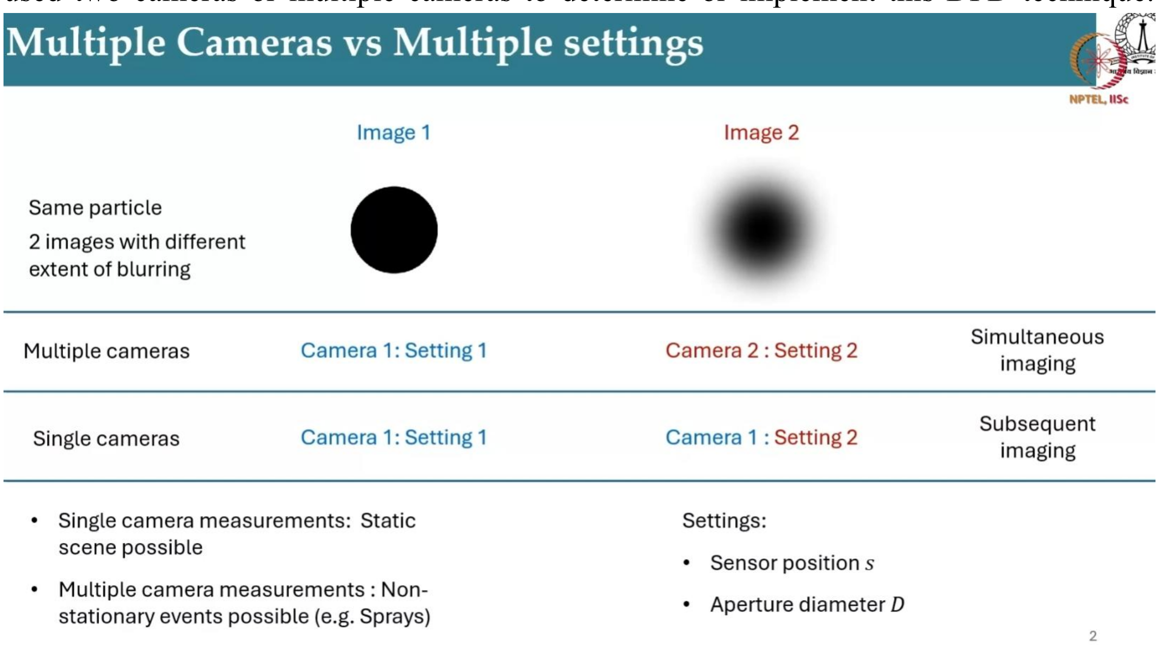
Prof. Saptarshi Basu
Department of Mechanical Engineering
Indian Institute of Science, Bengaluru

Week – 06

Lecture - 30

Depth from Defocus – 4

Hello, everyone, and welcome back. So in the last lecture, we discussed the convolution integral equation, which gave us an analytical equation of the blurred image formation. Apart from that, we understood the concept of thresholding and how it is useful in determining the apparent particle size, and we used that idea to develop a technique known as the two-camera depth from defocus. Where we simultaneously capture two images and determine the particle's position and size. So, in this lecture, we will see some other approaches to implementing depth from defocus and what some of the drawbacks or advantages of this technique are. So first of all, as we have seen in the last lecture, we used two cameras or multiple cameras to determine or implement this DFD technique.

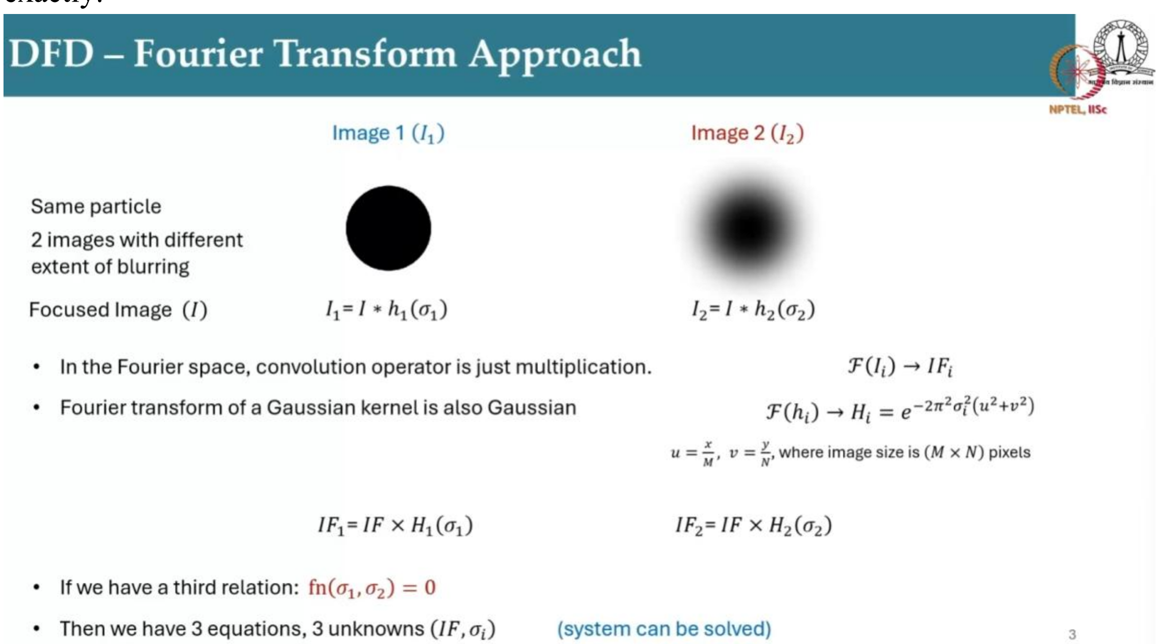


But can we do it in some other way? So what we have realized is that we need to determine two parameters. So, in some general sense, we need two parameters that can be measured, and then we implement two equations to solve the system in a complete sense to determine the depth and size of the particles. But with a single camera, it seems that it is difficult. But there is one more approach.

What if we use multiple settings on the same camera? So, the earlier configuration worked with multiple cameras with different settings. So, camera 1, setting 1; camera 2, setting 2. We were doing simultaneous imaging, but in the single camera case, what we can do is have camera one with the first setting and the same camera with a second setting. We then have to take subsequent images. What I mean by that is you take the first image with the first setting, then change that setting, and then take the second image.

What are the consequences of such a measurement that your scene has to be static or stationary? You cannot have moving entities in such a scene, right? Only then can you do such kind of subsequent imaging, but with multiple cameras, you can also work with non-stationary events, even high-speed events such as sprays. That's why in most of the implementations, we will have multiple cameras or things like that. But just for the sake of completeness, I will briefly discuss this multiple setting as well. And by setting, we mean either the sensor position, the aperture diameter, or other things, like in the case of multiple cameras, where the setting is associated with the multiple sensor positions. So the path length of the light or the sensor position related to lens S was being changed.

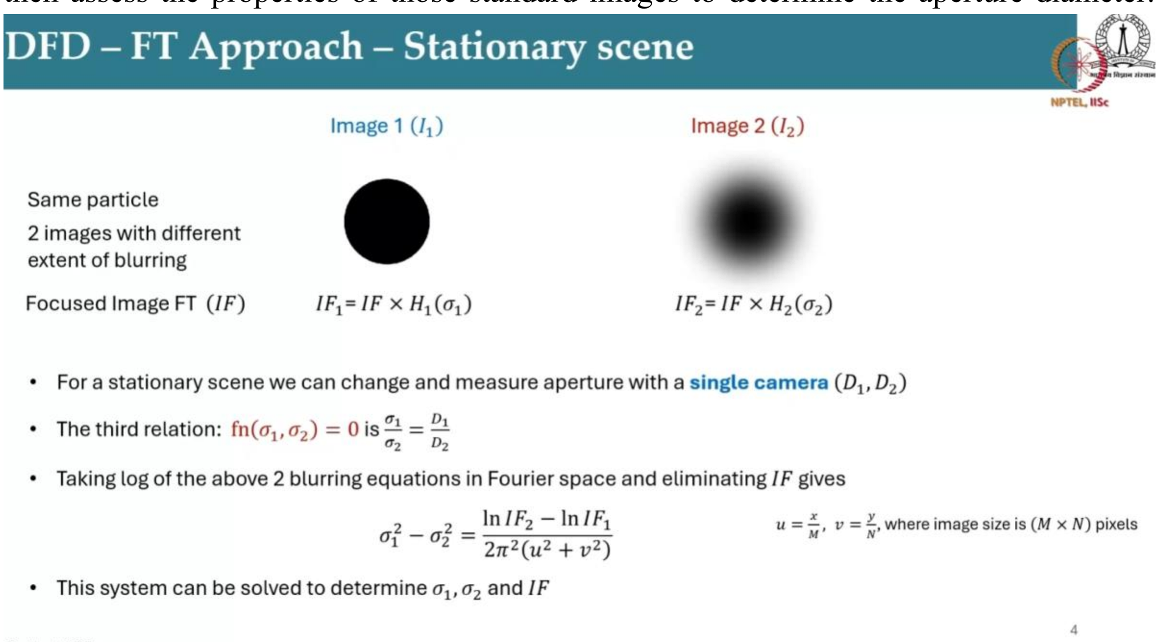
And with that respect, we were getting a different extent of blurring. But there are other approaches to do that as well. We will now talk about the Fourier transform approach and its implementation in both single camera multiple settings and multiple camera multiple settings approaches. So, first of all, this focused image, if it is represented by i, This image can be represented in this form where the focused image is blurred using a blur kernel; the first blur kernel, and the second image is obtained by convolving it with respect to a second blur kernel, and these are the two unknowns. We don't know it exactly.



Why are we talking about Fourier transforms? Because the convolution operation is very simple in Fourier space. It is just multiplication. So, first of all, if we apply the Fourier transform to the image i, then it seems we are representing it as if the Fourier transform of the Gaussian blur kernel is also Gaussian. So, it looks something like this: instead of x and y as the position, we have u and v, which are normalized by the total number of pixels in the x and y directions. So, it is basically in the range of 0 -> 1; nothing complicated here.

And if we apply the Fourier transform to this convolution equation, we get something like this, where the convolution operation is just converted into a product or multiplication. Now we have two equations. And if we have a third relation correlating σ 1 with σ 2, we have three equations and three unknowns. The three equations are this convolution equation for the first image and for the second image, and this correlation between the blur kernels and the three unknowns are the actual focused image and the two extents of blurring, Σ 1 and Σ 2. So ideally, this system should be solved, but is it solvable? Let's check that out.

If we use only a single camera with multiple settings, as I have discussed earlier, for a stationary scene. This works only for a stationary scene. Then you can perhaps work it out. How do you work it out? What is the setting that you chose? So the setting that we chose is the aperture diameter. So actually, you can measure the aperture diameter in certain cameras, and there are specific ways in which you can take standard images and then assess the properties of those standard images to determine the aperture diameter.



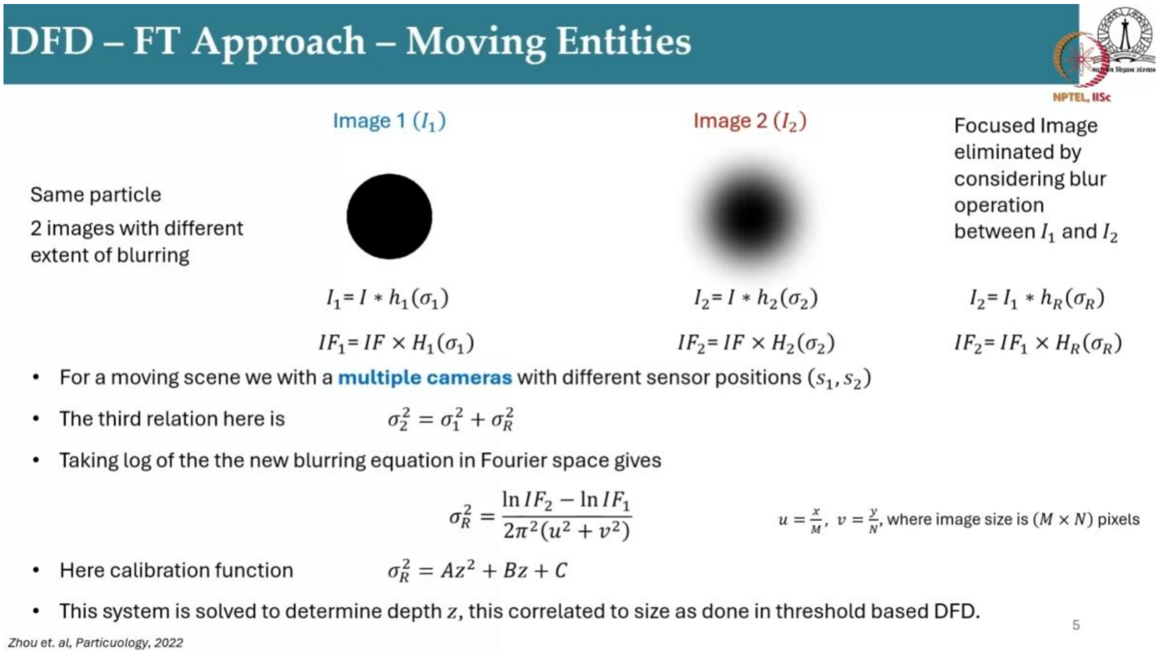
But controlling the aperture diameter in most cameras is very easy. Now let's say we know the aperture diameters D_1 and D_2 corresponding to setting 1 and setting 2 using the same camera. Then we have already determined that the blur circle diameter is proportional to the aperture diameter, and the standard deviation of the Gaussian blur kernel is directly proportional to the blur circle diameter. This gives us the third relation that we were talking about in the last slide: the blur kernel diameter is proportional to your aperture diameter, and we have this equation for the two images. So now we have three equations.

This equation represents the blurring in image 1 and image 2, and this equation correlates the extent of blurring of image 1 and image 2 because of the knowledge of the aperture diameter. Taking the logarithm of these two equations because we know that H has an exponential form and simplifying. By simplifying, I mean we can directly eliminate "if." If we take the log and divide equation 1 by equation 2, we will get something like this by rearranging and other similar operations. Here, we clearly have ln.

So, we can apply the Fourier transform to the actual captured image and take the log of it for both images, and u and v just represent the positions x and y, while σ_1 and σ_2 are associated with the extent of blurring for image 1 and image 2. So, it is very curious to observe that you can take any particular x and y, and this left-hand side should come out to be the same because this is a position, right? You have, say, an m x n image, so you have m x n number of points; you can choose any point, and it should come out to be the same value. Obviously, there is noise in your imaging system, so it will not be exactly the same, but usually we should choose an area where there are a lot of details that are blurred out. Therefore, applying this equation is the most reasonably good option, or you can apply it to all the points and take a mean or something like that, whichever way you prefer. So we have the value of $\sigma_1^2 - \sigma_2^2$, some value, and we also have $\frac{\sigma 1}{\sigma 2}$, some value.

Okay, so we have two equations and two variables now because we have eliminated, if so, this will give you directly the value of σ_1 and σ_2. Now you can substitute σ_1 and $\sigma_2$back in either image 1 equation or image 2 equation. Determine the actual image of which you can then take the inverse Fourier transform to determine the actual focused image, so in that way we can apply this technique for the case where things are not moving, and you can take subsequent images by changing the aperture diameter. For sprays, I agree this is not. At all practical, but for other scenes where you have some dispersed particles, lay say dispersed or fixed over a volume, just for example, I can think of these acrylic blocks.

If you cure an acrylic block or some polymer block and you don't remove the air from it, those tiny bubbles will be dispersed throughout, and you can then measure the sizes of these bubbles. Also, the spatial distribution of these bubbles using this technique where they are not moving. It's a static scene. Coming to the Fourier transform approach for moving entities, first of all, we cannot take subsequent images.



So that has gone. So we have to use multiple cameras in that case. And in this case also, we implement two different sensor positions that we discussed in the previous lecture. Like we have S_1 and S_2 for camera 1 and camera 2, correspondingly, in the same way, as I expressed in the last lecture, you can apply the Fourier transform to the blurring equation, and you get two equations directly, which will remain the same in this case as well. Apart from that, one interesting thing that we are trying to do here is that we can directly eliminate the focused image by considering a blur operation on image 1 to obtain image 2. So what we are trying to say is, what if I blur? Suppose we establish our imaging system in a way that image 1 is less blurred and image 2 is more blurred; then we can apply blurring convolution over this image 1 to get image 2, right? So there will be a blur kernel.

σ_r, which, if applied to this image, will give this image, right? So we can write $i_2 = i_1$ convoluted with the blur kernel with some result, some σ_r. Okay, and if we apply the Fourier transform, we get an equation like this: how are σ_1 and σ_2, and σ_r correlated? They are related like this, so the final blur kernel σ_2^2 of this image is equal to $\sigma_1^2 + \sigma_r^2$. Which holds true for any systems with standard deviation, so if you aggregate two systems with standard deviations σ_1 and σ_2, then the final result will be $\sigma_1^2 + \sigma_2^2$. In this case, we are aggregating $\sigma_1^2$with $\sigma_r^2$to get σ_2^2, right? So this is the third relation that we are trying to state, but it's still not in. It's still not complete; it's not self-sufficient to determine because we still don't know what σ_r is.

Right, we don't know σ_1, we don't know σ_2, and we don't even know σ_r, right? How to do that? So in this equation, this new equation that we have found out, if we take the log and rearrange it, we get σ r squared to be something like this. So here again we have IF_2, we have IF_1, we have AT, we have the left-hand side basically. So if we have the left-hand side, we can easily estimate σ_r^2, right? But just the knowledge of σ_r^2 is not enough. We know how image 1 correlates to image 2, but we don't know what the depth is by which both of them are. Displaced, or what is the depth position of the particle, something like that? So, we have to generate a calibration function now, like we did earlier in the previous technique with two cameras.

We correlate this σ_r^2 with depth z, and this functional form is something that can be deduced from geometric optics. I am not going into the detail; you can refer to the citation here. But it's fairly straightforward. It is very similar to what we did while determining the blur circle diameter. So you can clearly show that σ squared does have this quadratic functional form with respect to position.

All we have to do now is determine the coefficients a, b, and c from experiments that can be done by taking target dots at known depth locations, and at each location, you determine the σ_r^2. And if you plot that σ_r^2 with respect to Z, you will get a curve. You can fit the curve in the region, and you can determine these values A, B, and C. In that way, whichever image is captured, you have the σ_r^2 value of that. Then you can solve this equation to determine Z.

So depth is easily estimable using this Fourier transform approach. But apart from that, you can also determine the size. If you also take the thresholded size during the calibration in the same way that we did in the two-camera approach, you will have this depth versus size correlation, so in that sense, you can correlate this σ_r^2 with size as well. So it all depends on what you want to measure. If you just want to measure position, you don't need to measure the thresholded sizes at each target dot location.

DFD – FT Approach – Moving Entities

NPTEL, IISc

- Calibration function $\sigma_R^2 = Az^2 + Bz + C$

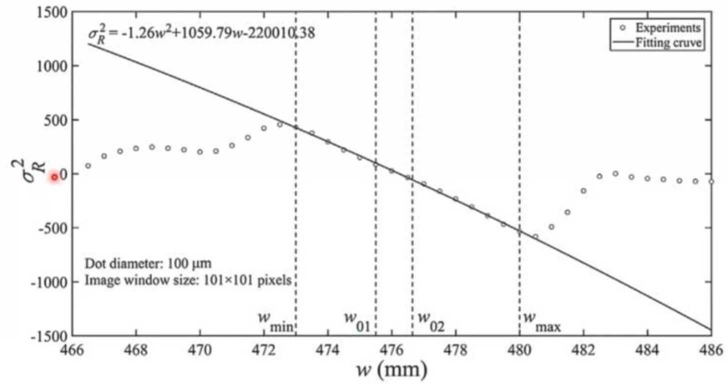


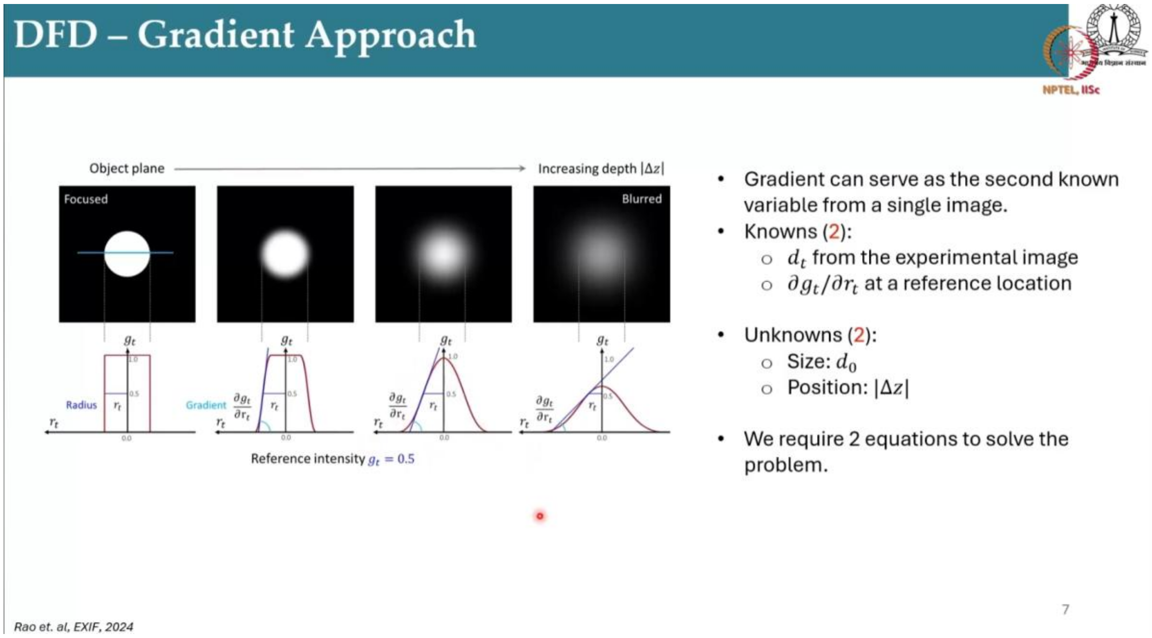
Fig. 5. The calibration curve and the depth measurement range.

6

Zhou et. al, Particuology, 2022

But if you want to estimate depth, then you can correlate the depth of the target dot, the known depth of the target dot, with the σ_r^2 to generate a calibration function. So, the experimental calibration curves look something like this. Here you will see that at each depth position, just represented by w, the σ_r^2 values are plotted, and you will see that this kind of quadratic functional form is valid only in this region, and this region then demarcates your measurement volume. To the point to which this function is valid, it demarcates your measurement volume or detection volume, so in this case, you will also have prior knowledge of your detection volume. Now we have discussed depth from defocus based on thresholded diameters and based on the Fourier transform of images.

The third approach that we are going to discuss is dependent on the gradient. And the gradient, I will just demonstrate it using this schematic here. What we have done is provide these blurred particle images. So, here it is in focus; here it is extremely blurred. So, what we did was take a radial slice and plot the intensity across it.

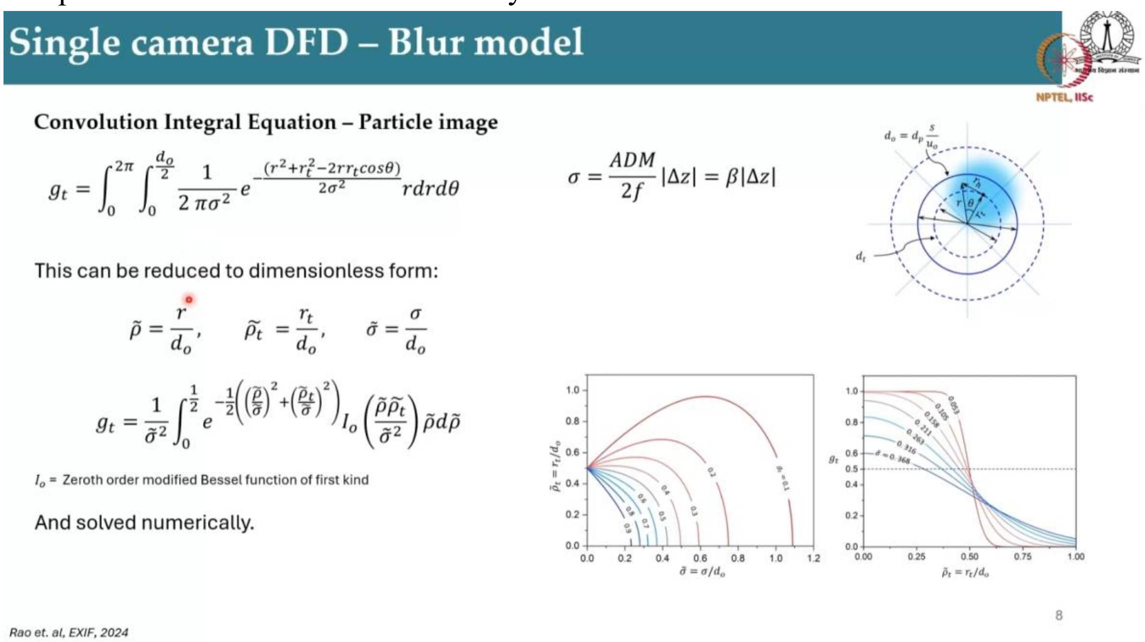


So, for the sharp or focused images, you will see a step function kind of thing, a top head kind of function, right? There is a step jump, then a constant, and then again a step jump. Here, determining the radius is pretty easy, but as you blur it, you will see that there are smoother jumps now; like, there is this smoothing around the edges, right? Some kind of variation occurs over a larger distance. Here, the variation is happening at one pixel level; okay, here it is happening over a few pixels, and here you will see the variation is even more spread out, or the blurring is spread out. For even more blurred images, you will see it is extremely spread out. So we realized earlier that this thresholded radius r_t, as plotted here for threshold intensity 0.5, is changing; right here it is the same as the particle size, here it is slightly smaller, here it is even smaller than the actual particle size, and here this r_t is very small. You can clearly see that after a certain point this dome will be below 0.5 and the particle will not be detected because there are no pixels in your image where.

.. Uh, where the threshold intensity is, that logic is obeyed. Okay, but apart from that, you will observe one more interesting thing: you will see that at that reference intensity of, say, 0.5, if you estimate the slope of this variation, you will see here it is just 90^0, right? It's just a vertical line, but here you will see it is slightly lower; the gradients are slightly lower. Here this gradient is even lower, and here it is even lower, which is what we intuitively expect as well because if the features are getting smoother, then definitely there are no sharp edges or things like that, and if we have smooth features, the associated gradients will be lower. It gives us a hint that the gradient can serve as a second known variable from a single image; earlier, we were only using this thresholded radius from a single image, and for the second input, we had to rely on a second camera or second setting to get a different thresholded radius or diameter that serves as two inputs.

Then, we had to generate two equations and solve them to get two outputs: d naught and z. But here, if from the single image we can estimate the thresholded radius or diameter, as well as the gradient at a certain intensity location, then we have two knowns from the image and two unknowns. To solve this system, we definitely need a minimum of two equations. That is what we have been seeing in the first few lectures. How do we do that? How do we generate those equations? We will go back directly to the convolution integral equation that we generated for the particle image.

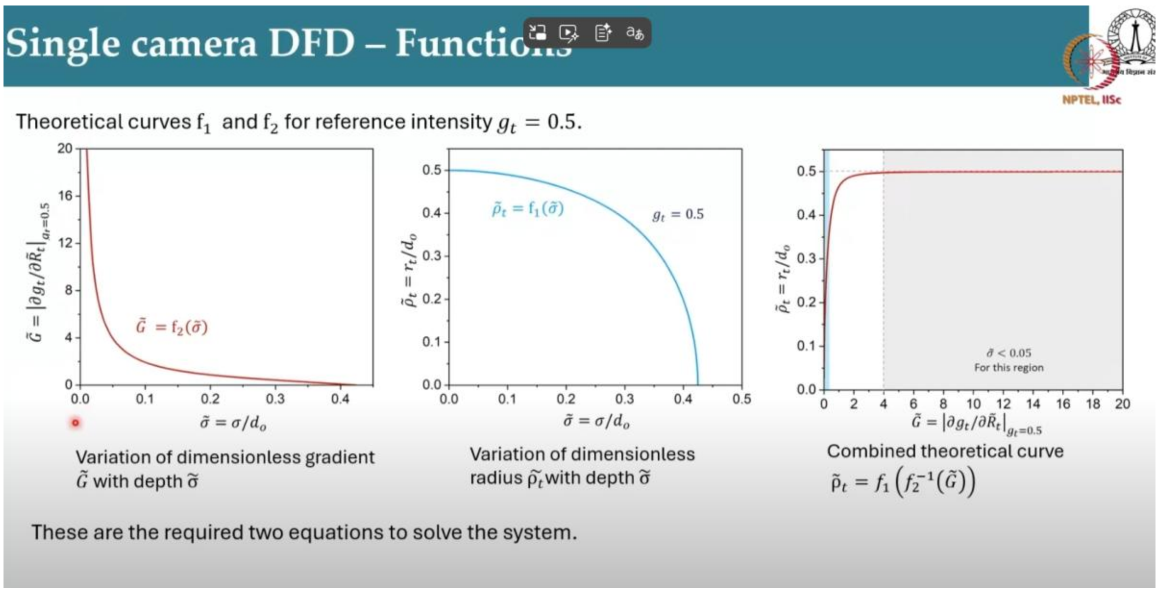
So this is the equation that we generated earlier. The only thing that we do before solving it is to normalize everything with respect to the actual particle size. So we normalize all the integral variables with respect to that and σ also with respect to that, and as this is theta symmetric, we can integrate it with respect to theta, and we get this. So this is the final non-dimensional form of the convolution integral equation, and we solve this numerically, and these are the solutions. So this solution is representative of the depth compared to your thresholded radius.



So if we choose a threshold of 0.5, we will look at this curve, this blue curve. Okay. So what it tells us is that with increasing blur, the thresholded radius will decrease and eventually go to zero. At a certain degree of blurring, and this is what we were talking about earlier as well, and if you choose a different threshold, you can detect it for, say, larger degrees of blur, and if you choose a higher threshold, you will detect it for even lower degrees of blur. So this x-axis is your degree of blur, and your y-axis is the thresholded radius, and as we know, σ is directly proportional to z, so this x-axis is actually representative of your.

Depth parameter, and this is nothing but the radial distribution of intensity because your y-axis is intensity and the x-axis is the non-dimensional thresholded radius position. So at each radial position, basically, you will estimate the gradient, and you will see at a lower σ value. So these are σ values; all σ tilde are basically normalized blur kernel values. So if at a lower σ, you will have sharper features and higher gradients. At a lower σ, you have smoother features and smaller gradients.

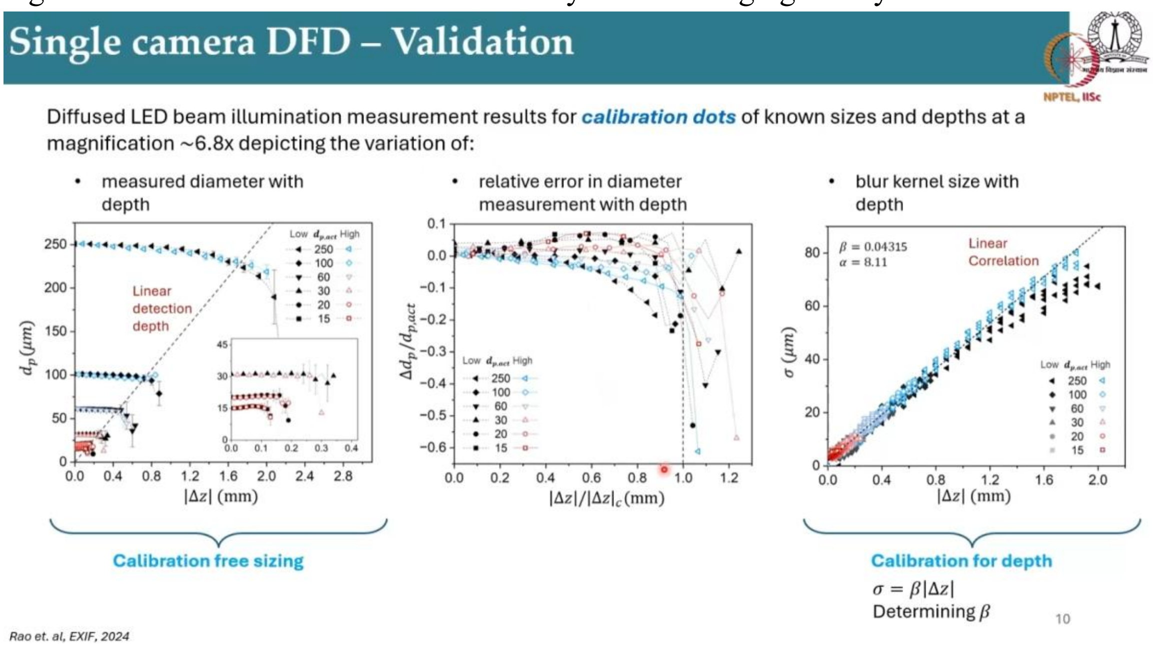
So these are the solutions that we obtained from this equation, and we will use them in this gradient-based single-camera depth from defocus. So from the previous curves, as I have shown you, the gradients were varying. We can actually estimate the gradients at the reference intensity location of 0. 5.



We chose the midplane for this problem. You can also choose other values, but we found that this is the most suitable one. We can correlate this σ with the gradient or the extent of blurring with the gradients, and you can see that when the extent of blurring approaches 0, the gradient approaches infinity, which is like a very sharp edge. And we also have the correlation of the extent of blurring or depth with the thresholded diameter, and these are the two equations that we need to solve the system. If we combine both of them, we get this curve, so I will not be discussing this here. But, yeah, these are the two different equations that we get along with the two knowns that we have determined: the thresholded size and the size.

Threshold size and your gradient from the image are important. Then you have these two equations, and using this, you can determine the extent of blurring and the actual particle size for the system. Now, for validation, we started with calibration target dots. So we took calibration target dots of known sizes and moved them in the measurement volume at specified intervals as if we knew the distances. For the z positions for each of them, when we applied this technique, first of all, just to determine the size of the particle in the way I have expressed everything till this point, no calibration curves were required; we just used two theoretical curves, which gave us the two equations, and using just that, we can determine the particle sizes.

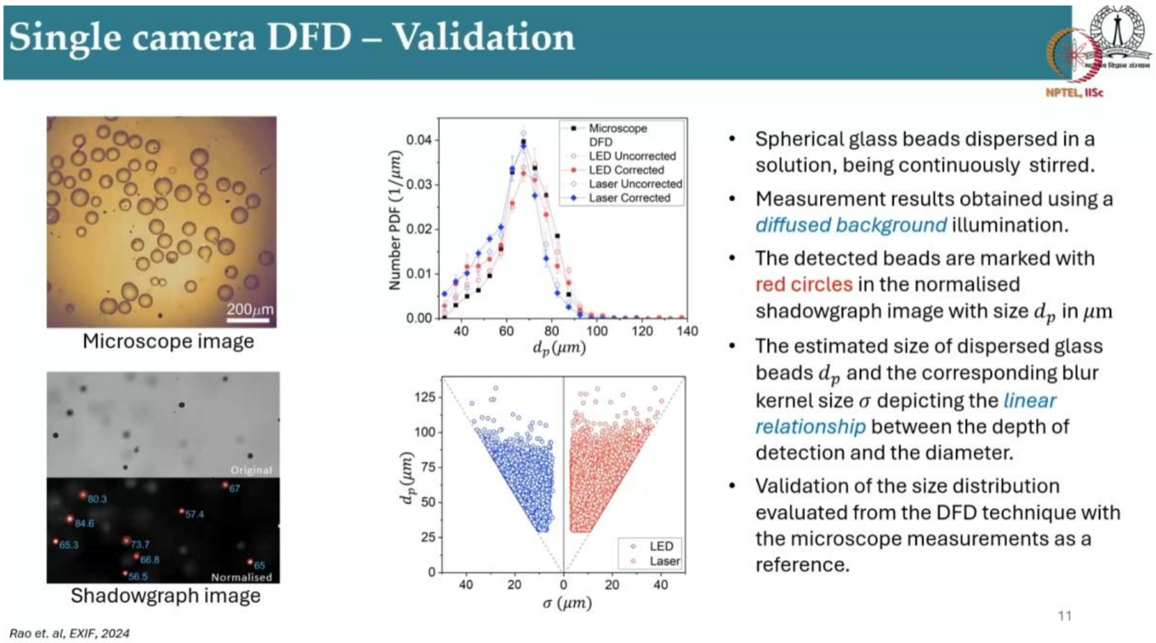
It looks something like this. So here for the 250-micron particle are the measurements at different depth locations, and for all the others, it is shown here. You can see that they are pretty accurate, and this part, as I expressed, is a calibration-free sizing. Here is the relative error in diameter measurement, and you can see it is pretty accurate within plus or minus five percent to significant depth locations. And this dotted line actually marks your depth of detection, which we are imposing artificially, because beyond it, the errors are pretty high due to effects like very low intensity values, as the blurring is extremely high and the noise in your imaging system dominates.



The second part is the estimation of depth. So we know the depth for each particle, and in this process of calibration-free sizing, we estimate the σ. So if we plot that σ with respect to the Z position, we see a straight line, and this is the linear correlation that we also expect because we know Σ is equal to beta Delta Z, and beta is a constant for your imaging system. We need to determine this beta, which can be done by just straight line fitting or linear regression over this data. So what this brings us to is that if you want to implement this in a spray, you will directly get something like this: you will get all the sizes, but you will not have their spatial positions; however, you will have the Σ values for each of them. So for each of the droplets, you will measure D naught and Σ.

To estimate delta z, you have to do this calibration procedure for your imaging system. So you have to take target dots, move them over a specified volume at known depth locations, estimate σ correlated with delta z, get the beta coefficient, and then apply or implement those beta coefficients directly over the spray measurements as well because they are constants for your system. So it will remain the same, even for your spray system. So we applied it over some spherical particles as well because the target dots are nothing but 2D discs. So we took spherical glass beads, placed them over a slide, and imaged them under a microscope.

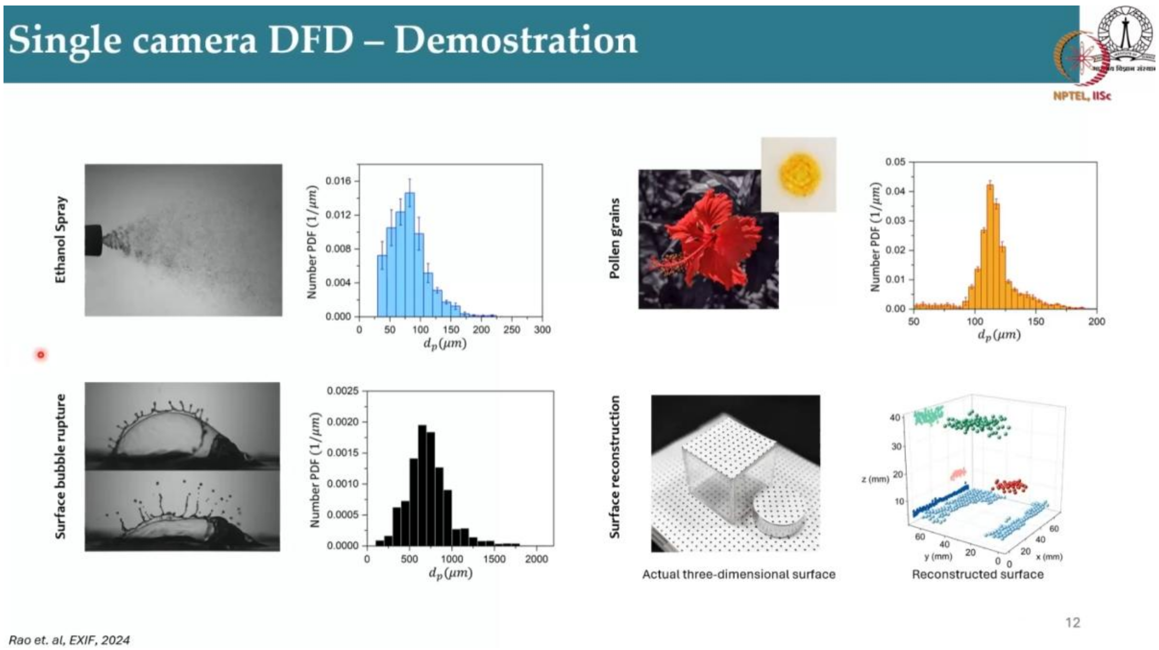
Through that, we obtained the ground truth distribution. So this black curve actually defines the distribution present. In this glass bead system, it is measured in a dispersed fashion, where some of them are in focus and some of them are not. We put them in a liquid solution that was being stirred continuously. So, this is one of the snapshots. You can see that some of them are in focus and some of them are not.



Then we normalized the image and applied our technique to determine the size. So, these are the sizes. And if we plot a probability distribution function for that, we get something that matches pretty well with the black curve, right? So the blue and red ones are related to the different kinds of background illumination, which also affects the results, but I won't be discussing them in depth right now. But you can see that it agrees pretty well with the microscope data. And also, if we plot the diameter versus the σ, which is representative of the depth, we see that everything is enclosed within a cone.

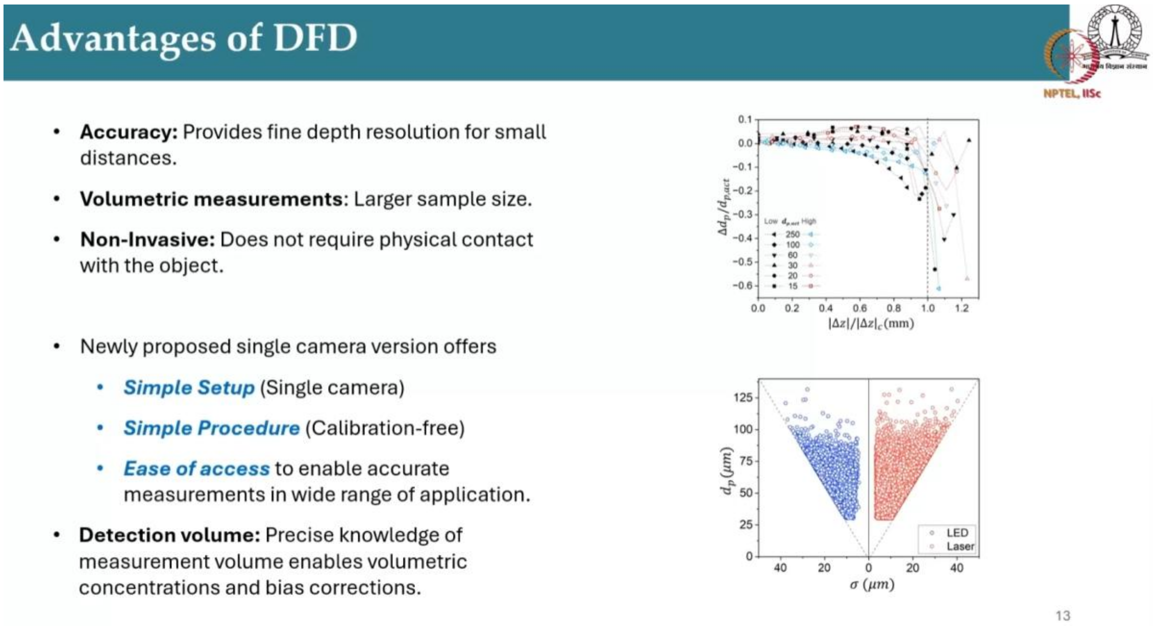
And this linear detection edge is something that we discussed earlier, even in the two-camera methods. So it's the same thing here: the particles of smaller size get blurred easier and are measured over a smaller volume, while the larger particles are detected over larger volumes. So the same thing is valid here as well, and we will use that information in this case for the PDF correction or bias correction, which is represented in this plot as well. So we have plotted both the corrected and uncorrected estimates, and you can see that it works pretty well.

Then, these are more test cases where this technique was applied. In the first case, we are measuring an ethanol spray, and this is the probability distribution function obtained from it. We even applied it to pollen grains, where small spherical pollen grains were dispersed in the liquid solution, the same with glass beads, and we get this distribution. We can even measure droplet sizes generated during a bubble rupture. You can see that in the surface bubble ruptures, the ligaments break down and form this array of droplets, and as this is a three-dimensional phenomenon, it kind of flings droplets in and out of the plane of focus. That's why, because the number of points is very low, sampling the droplets in the whole volume then becomes very important to get the whole sample.



So, the DFD kind of technique is very... It is suitable for implementation in such cases; if we calibrate σ to delta z, then we can convert the σ values to delta z for 3D depth reconstruction as well. We have printed these target dots over a three-dimensional surface, took only a single image of that, and from that, we were able to reconstruct this surface with an accuracy of five to ten percent, which is pretty good because we are just using a single camera, which makes it really effective. Easy to implement. So you can use it in systems where it is not that easily accessible to take measurements.

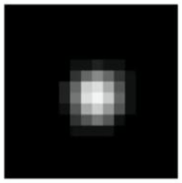
So now I will quote some of the advantages of the technique. It provides reasonably good accuracy, considering the simplicity of the setup. It provides volumetric measurements; the sampler size is larger compared to the other available optical techniques, where you do point measurements or planar measurements. Here, you measure the dispersion within a volume, which definitely improves your statistics. This method is non-invasive, and since it is an optical method, all you need is optical access to your system, making it very useful in a lot of scenarios. The particularly, if I talk about the last single camera depth from defocus, is a pretty new technique, and it's fairly easy to use because there is only a single camera, and the procedure is also very simple because it is calibration-free.



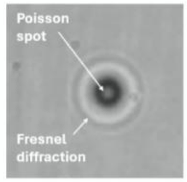
So if you want to just measure the sizes, it is calibration-free, right? And in that case, it allows for ease of access. Because it just requires a single camera and a diffused background light source, you can easily implement it in systems where access is not that easy. So, if you cannot install big lasers and things like that in remote locations, then you can use this technique. One of the most important advantages of this depth from defocus is that we have precise knowledge of the detection volume, which is crucial for volumetric corrections, concentration determination, and bias corrections, because in other techniques, such as point measurements and planar measurements, they are also volumetric measurements in some sense, but the volume is confined to very narrow regions; in point measurements, the volume is very small, so it is kind of a point measurement, and in planar measurements, the volume is very thin, so it is still a volumetric measurement but with a very thin volume. These bias corrections are not easy and are very tedious to do, but in this case, it is pretty straightforward because we can determine this bias theoretically, which makes the technique very powerful.

The limitations of this technique include, first of all, capturing discrete information when we use any practical imaging system or camera. So the images are composed of pixels, and we have discussed that they look like a matrix. Estimating gradients from such information where there is also noise is really challenging because gradients are very sensitive to noise and averaging effects. So that's why some inaccuracies persist in the gradient-based DFD that we talked about; even the thresholded diameters are sensitive to that.

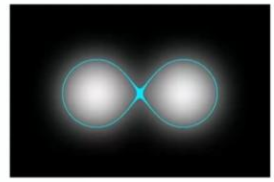
Limitations of DFD



Discrete information:
Image composed of pixels, restricting the precise estimation of gradients at exactly the reference intensity.



Non-Gaussian blurring:
Fresnel diffraction and Poisson spot observed in collimated background illumination.



Concentration Limit:
Particle overlapping can result in a single indistinguishable, non-symmetric entity due to blurring.

Rao et. al, EXIF, 2024

14

But somehow, using interpolation, we can get rid of it. However, the gradients are more sensitive to it than the non-Gaussian blurring. All the things that we have discussed until now include some other PSFs, but we stuck to the Gaussian PSF. But if you have background lighting or backlighting that is coherent in nature, then you will have these diffraction effects as well. The area disc that we talked about is the limiting PSF in the case of diffraction effects when the image is in focus, but when it becomes defocused in a setting where the diffraction effect is superimposed over the defocused effects. Then you will find this formation of rings; you can see this is the blurred particle image.

At the center, you see a bright spot, which is known as a poison spot, but around the particle, you see these rings. This is because of rational diffraction. So this is very difficult to deal with. Here, you don't have a monotonic increase or decrease in intensity in the radial direction.

You have non-monotonic changes, and gradients are very sensitive to that. So if you try to implement the gradient-based technique on such an image, it will definitely give extremely wrong results. Maybe in the threshold-based two-camera DFD, where you are calibrating your system experimentally, there is a chance that the calibration curve itself will adjust to these effects because you are doing it for that system. But still, it can give you inaccurate results in the limit of very small particles. The third limitation is the concentration limit. If you remember, when we started this part of the course, I showed you some images of Like dense sprays, right? So then, in that case, you have a lot of overlapping spray droplets, and in such a case, all the techniques that we have discussed are very difficult to implement.

We still are in the stage of developing new techniques for such scenarios. To be precise, even the other non-imaging based techniques do not work on such dense spray systems; there is always a limitation in the number concentration of your dispersed entities in the volume. So all these poses a very like challenge to implement DFD in the way that we discussed like having the precise knowledge of the exact PSF, getting rid of the noise on your system or the getting around the overlapping effects is something that people can look into in future. With this, we come to the end of this part of the course where we discussed a very fascinating technique known as depth from defocus. In this, we correlated depth with defocus, or the amount of blurring, and we used that blurring information to determine the size and position of particles or droplets.

We did it in different ways. We used it; we determined it using gradients. We determined it using the Fourier transform. We relied on a thresholding approach to determine that. So, there are various approaches that can be implemented, but the core idea still remains the same. And there has been significant progress in this field in the past few years, special thanks to Professor Kamantopia, because he has been a significant part of these recent advancements.

You can refer to all these papers from which these ideas were taken. And with all the advantages and disadvantages that we discussed in the last few slides, we hope that this technique progresses even further, especially to eliminate those disadvantages. Thank you.