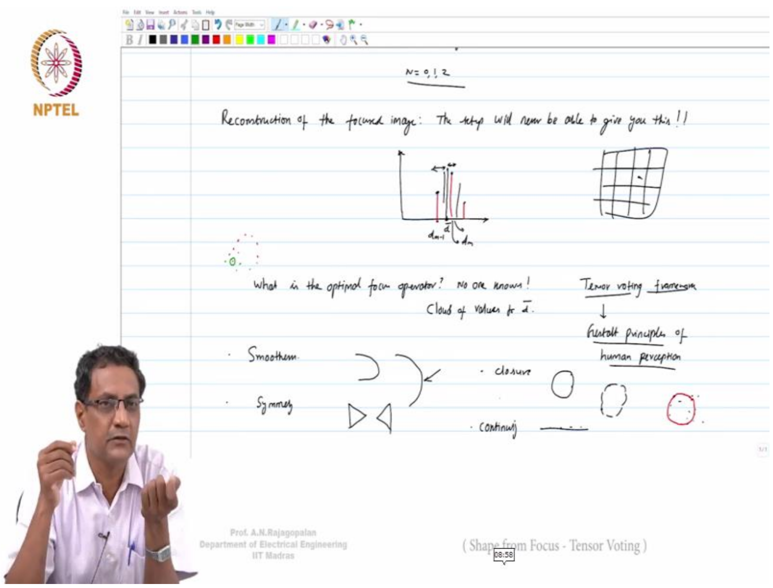
Image Single Processing
Professor. A. N. Rajagopalan
Department of Electrical Engineering
Indian Institute of Technology, Madras
Lecture No. 33
Shape from Focus - Tensor Voting

(Refer Slide Time: 00:16)



What is it what is an optimal what is the optimal focus operator? Nobody knows. What is the optimal focus operator no one knows, no one knows but then what I mean? If you if you ask yourself you might say that I should still have a way I should still have some way by which by which actually I should be I should be able to use all these all these operators together is, is there a possibility to, to kind of use it correctly you can always ask that thing.

See now it looks like you just apply an operator you get some D bar user operator everywhere same operator and then you have a depth map but then since we know that there is not just one operator, there is so many of them. And also even, let us say within each operator there arethere is this kind of a variation.

For example, if you look at some modified Laplacian, we could have taken an average, like by taking N equal to 0 that means, we simply we simply look at the modified Laplacian value at that pixel, or we could have taken N equal to 1 which means the sum of the values and a kind of a 3 cross 3 neighborhood, we could have taken N equal to 5, then we will get a 5 cross 5 neighborhood and you will notice that for each one of them, you will get no he will get a value of D bar, which could be changing.

For example, if I take N equal to 0 it means it might say that right your ACD bar is this if I take if I take if I take an average of 3 cross 3 values of ML, it may say that your D bar is no this, we take a 5 cross 5 it may say that your C D bar is something else, which one of these should I pick? I do not know a prior, which one of these is correct. And then again if I go to another operator, it might say that your D bar is that.

Now it could be that it could be that they will not be too far away from each other but there are these multiple, multiple values that will show up if you try to use different operators, or even the same operator apply that applied with, let us say varying sizes and so on. So, you can see that you almost can, can get a sort of, a cloud of cloud of values for a D bar.

For at each pixel, by the way, I am talking about cloud of values for, for a D bar, that there is a framework called which is actually a Tensor Voting Framework. For those of you who are interested, you can go read about it, there is something called a Tensor Voting framework, this is not this is not anything related to this problem or anything. Tensor Voting has been around for a long time.

This kind of a voting framework, this is based upon what are called what is called a Gestalt, Gestalt principles of human perception. So, so this so this kind of Tensor Voting framework is something that actually, that actually mimic, mimic this, human perception. What that actually means so now there are about 6-7 of these principles. Some of them I will tell you so, one of them is like is like what is called what is called a Smoothness which I said, which is some, which is like a generic prior.

So for example, so for example, if you read them no, no, no, no if you say that should I should I should I choose a curve like this over a curve like that I would rather, rather go for something like that which is more smooth in terms of curvature here you have a sharper curvature here you have a, you have a kind of a reduced curvature.

So, it will try to typically go for, go for go for something that is kind of locally more smooth. Then there is something called, called a Continuity or Symmetry, All right, the next is something like Symmetry. So, what this means is that I mean, if I have read the when I think of think of something like this, if I have on this side and on this side, I have this, then, then humans typically would, would tend to think that should have been something like that.

So, the Symmetry is something that that we sort of inherently believe that believe it exists in the world. I mean, I say, it does not mean that you blindly enforce this and Tensor Voting is smart, it does not mean that it will directly go and enforce although they look for these things, it is a framework that will sort of account that will try to incorporate all these factors into it. So, so, so, so, symmetry is something like that, if you see read if you see something read in an image then, then if there is a partial information about the same object elsewhere, they would want to believe that probably this the whole object exists there.

So, it is like some kind of a symmetry that you expect then what else is there then there is there is a Closure, Closure means that I mean If I gave you gave you something like that I mean, you would you would want to get us to close this. But you would think that, that is probably this object should have been. That is a way our human mind works, we tend to, we tend to look for things that are democracy pleasing to the eye or that we feel that we feel should ideally probably exist in this world.

So, we tend to kind of close or for example, if I, if I gave you something like this then, then you would think that all in all in between something should have existed probably you missed it at and you, and you, you try to kind to see close all those gaps. Closure is one thing, then there is there is a Continuity, Continuity basically means that, if I gave you, gave you a bunch of dots like that? You would you would think that, this probably was a line, again, that these are all things that we that we kind of self impose. We do not we do not get to explicitly say these things, but then implicitly we do this.

Somebody shows something then implicitly we are kind of filling in all the blanks. And these blanks we have we are filling in typically based upon things that we that, that are more or less rational symmetry, smoothness, closure, continuity, things like that. So, this, so, this Tensor Voting such framework and then one more thing, suppose I tell you it can even be it can be a can even be used for kind of what you call used, used to kind of remove outliers.

So, for example, if I gave you points like this, and suppose, I gave you a give you a few points like that, which are outliers. And then so, suppose I asked, you try to construct, construct a surface or construct a curve out this that is meaningful, then it is very likely that very likely that you as a human will, will kind of draw a circle through that and then, then ignore everything. I did not draw it correctly.

Where was this point? One minute, let me draw it properly. So what it will mean is, so when you see this, you will end up, kind of drawing a circle through these points, and you will say that, if I, if I had to, if I had to visualize an object through this, that would be the object and then say, rest of the points for me are all outliers. So that is like, noise, we should tend to ignore. Is it not true that we can attend to do these kind of things?

So, you will automatically knock off some of those points, because you are, again, implicitly looking for some sort of a structure there, which should be regular, and so on. So all these things, so this Tensor Voting framework is something that actually that can that can do all of this using the terms like what is called the Saliency map. So, it has something like a curve saliency, structure saliency, surface saliency, and so on.

So, the so, so, what have what what, what can be done this you can actually throw all these D bar values. So, for example, if you, if you say that see for example, again let me kind of let me tell you that you can you can also use it to close like I told you so, what can happen is if has something like that and for this point if there are multiple, multiple options available to me suppose, suppose let us say let us say one operator says that this should be the value another says that should be the value, another says this should be the value another says this should be the value.

Now, what will happen is this guy, this guy will look at the look at the local neighborhood and try to see which one of these is most meaningful to pick. So, in the sense that in the in this case it will say that property this is the one that that you should go with rather than this or this or this or this or that because that would be the one which will probably meet these be, these principles. So the locally smooth it does not introduce, introduce curvature that is that is very harsh.

And then you know, it looks like there is a, there is a symmetry going on. So, all that it will automatically incorporate to be able to, to be able to pick the one that it thinks is most suitable.

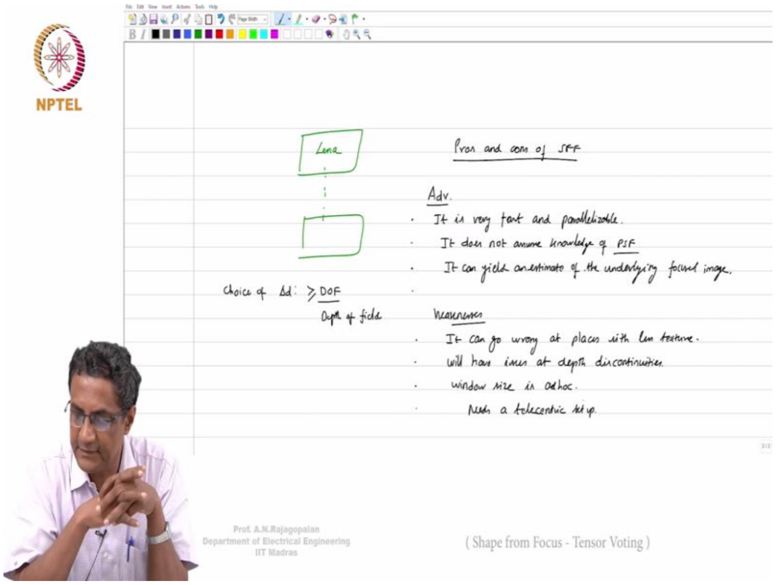
So, what has been done is? So, this so the Tensor Voting framework, I do not I do not have that paper. One, I am actually M-Tech students, he did this work. So, we actually showed that you do not have to stop with one operator, even we do not know which is the best operator.

But we said at least you have a we have a mechanism by which you can throw in multiple operators into the scheme and for every everything that we collect for say multiple, multiple values of D bar, and then use the neighborhood information and then do this kind of a Tensor Voting framework into in order to be able to pick the final value.

So all the I mean, we are not interested in going into the full details and all of that, the saying that do not think that, we are stuck with just one operator or something if you are willing to put in more work, then we can actually accommodate more operators. We do not have to, we do not have to shy away from that. We can take as many operators as we want. We can operate them at as at as many sizes as we want.

Come up with, come up with all these different votes for, for this one particular value and then see which, which one of them is most likely. So, basically this is all possible.

(Refer Slide Time: 09:06)



So, let me just talk about the pros and cons of SSF. So, the pros and cons. So, so the so, the first is the strength the this the strength is that it is, it is, it is very fast, it is very fast, I mean implemented in its fundamental form in the basic form, not with all this Tensor Voting and all that it is very fast when not with the parallax kicking in and all that. Very fast and a parallelizable, parallelizable.

So, it means that you can compute D bar independently of let us say other locations. You can run this entire operator, do the interpolation everywhere independent of the, of the other. So, if you have a GPU or something with multiple cores, then you could just run it all in parallel. Then two very important it does not assume, these are advantages of course, it does not assume knowledge of the PSF.

What is this stand for, a PSF? A Point Spread Function. So, it means that it is so, so, so we do not know where did we assume that the underlying blur is actually a Gaussian and so on. The one that we simulate for you is a Gaussian because that is what we can do at in the lab, but I am saying nowhere in these in this ring or in this whatever, that Baba face or that Lion thing, and nowhere did we assume anything.

So, all that all that we say is where does this point appear in focus? That is all we are interested in. We are interested in knowing what could be the Gaussian is it goes in is a non Gaussian we do not care. It is also that way it is it is a very good thing. So does not, it does not assume anything about the knowledge of the point spread function. Then 2 these two are obvious once it can yield an estimate of the focused image using of course, D bar estimate of the focused image or the underlying focused image, underlying focused image.

Then I think these are the main ones the ones that so, so this drawbacks and if you look at it the weaknesses, weaknesses are there. So, one is the one weaknesses, it can go wrong at places, at places with less texture which means that within an image wherever you find that there is not much activity it can go wrong because then all your D bar and all that you will not get a, get a nice Gaussian fit and then you will not, not know your know your D bar exactly.

Will have, a will have issues at depth adjusts or what is called what is called a discontinuity. Because if you are sitting right at the edge then your focus operator is like half of this is on this side, half of this is on the other side whether, whether the, the depth is suddenly falling. If it is a smoothly varying thing, then it is okay but is there is a sudden change in depth, then those locations will take a little while to recover.

Because somewhere I mean, somewhere you will be kind of half here and half there. And then the and then of course choice of del m, but that is not a weakness okay choice of delta F, I will just put this here, choice of delta D that somebody, this should be greater than DOF. There is not equal to okay at least this DOF is a Depth of Field. Then window size of operators is ad hoc, unless you use some like Tensor Voting again when I say all this is for the basic SF.

Window size is ad hoc is chosen in an ad hoc manner and then needs a telecentric setup again in the basic form. This, this should have come in fact, in fact in the beginning. So, some advantages some drawbacks and, and that is it and as far as, as far as this method is concerned.