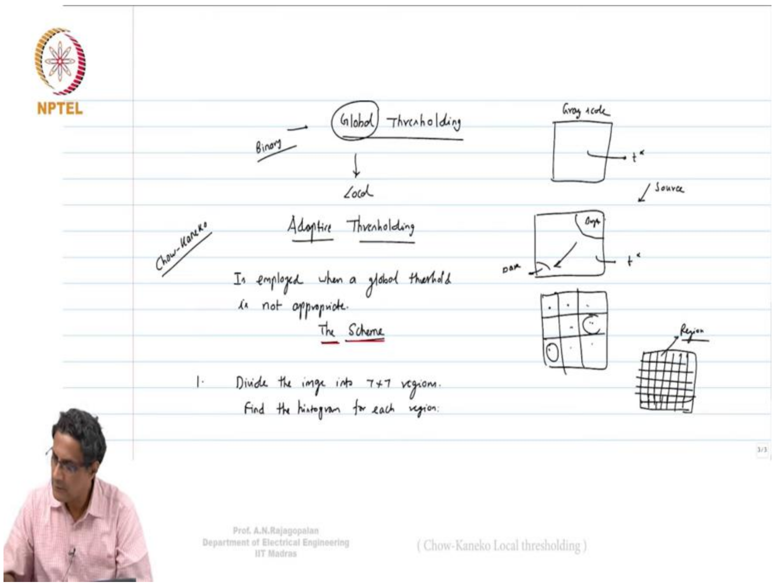
Image Signal Processing
Professor A.N. Rajagopalan
Department of Electrical Engineering
Indian Institute of Technology, Madras
Lecture 60
Chow-Kaneko Local Thresholding

(Refer Slide Time: 00:16)

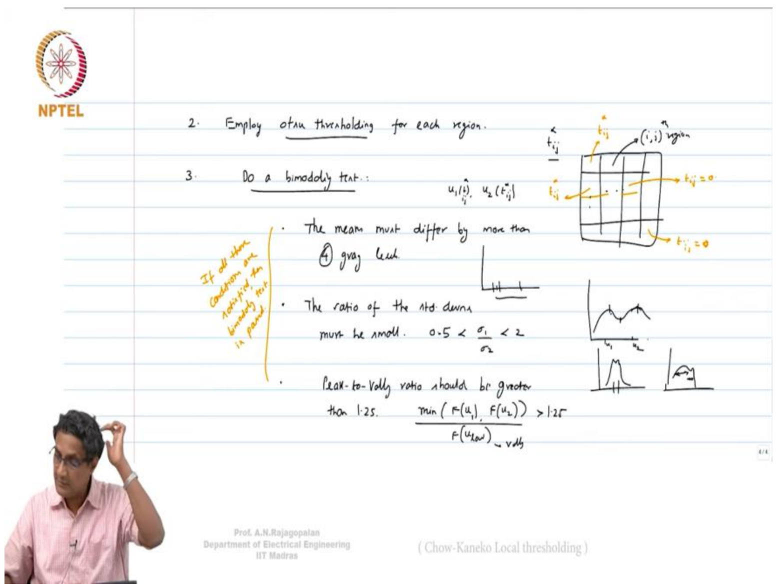


Now what is this scheme? So what is this scheme now? This adaptive kind of thresholding scheme, this scheme. This scheme. First thing - what it does is - now I am going to talk about something you know, that is borrowed from a paper by Chow-Kaneko okay. Now this is a thresholding scheme that comes from this paper of Chow-Kaneko but you know, but this threshold scheme is fairly general, this adaptive kind of a thresholding strategy - very general and you could apply it even for Otsu and so on.

But I am just, just for your reference. Now in this paper, right, what they suggest is - they say that divide the image into seven cross seven regions. So what this means is that, what this means is that I mean, so you had the whole image, right, and now you divide in seven cross - 1, 2, 3, 4, 5... 1, 2, 3, 4, 5, 6 and just one more box. And there you have like your seven cross seven, no each is a region - please remember this is not a pixel. This is a region.

Okay, each is a region so you have got like seven cross seven, there is not anything very sacred about seven cross seven, that is what they talk about in their paper. You can change it. You can make it five cross five, you can make it nine cross nine, it is up to you. But in the paper, right, what they talk about is seven cross seven. Now, find and find the histogram for each region. Find the histogram for each region, okay?

(Refer Slide Time: 02:17)



Then then employ, okay, in your case, in our case we are talking about Otsu so employ Otsu sort of a thresholding method - you could replace it with anything, okay. But since we have, since I have taught you Otsu, we will talk about Otsu. So employ Otsu at thresholding for each region. Okay, because you have the histogram, now you know how to actually search for a t. So for each region, you will search for a t star.

Now, okay, call it whatever t i j star. Let us all it as t i j star, okay, what do we mean by t i j star, so that is like in that region, okay whatever you have got like seven cross seven right - so whatever you have got like seven cross seven. Now any region is like the i j region, okay, a threshold for that - the optimal value of a threshold for that - we will indicate it as t i j star. Now after having found t i j star, we do what is called, we conduct, do a bimodality test.

What this kind of means is that because of the fact that I have actually broken it down into individual regions, we right, they advocate the idea that you should check for the fact whether the employing histogram can be reasonably believed to be kind of, right, a bimodal in the sense that whether the distribution of the intensities in each region whether it follows a bimodal or not.

Because after all, right, we are trying to, which is what we are trying to do in this case when you are trying to binarize - so it is saying that there should be actually two classes.

Now it could so happen that actually there is only one class under that region simply because you have divided the whole image, in which case, it would not be, it would be incorrect to try to choose some, say, t star or try to fir some, say, t star there. Okay, now what this bimodality test, what is it now? Okay, now the way it works - and I am going to write down the steps for this. The bimodality test, now the steps are as follows.

So what does the bimodality test involve? So first thing is - it says that the means, so after you have computed, right, so after you have found out a t i j star for a region, you find out their means and you would know the means, because this is what you have calculated like a mu 1 of t and mu 2 of t. Now in this case you will have mu 1 of t i j star and mu 2 of t i j star, okay. Now the means must differ by more than four gray level, by more than four gray levels.

Now there is again nothing sacred about this four. Like I said in one of the earlier, in my very first class on image enhancement I said that there will be certain hyper parameters that you might want to choose when you are doing image enhancement. So those some kind of hyper parameter. now you may want to choose six, someone else might say that I want to use eight, ten, okay, we do not know.

So what this means is that, right, so after you have divided it into kind of two classes, you are saying that these means should be sufficiently well apart. It otherwise, otherwise you know, it is not, so it will not be a good idea to have two means that are very close. Okay, and to sort of look a, say, t star okay, you know in between them. So which is why we write bimodal, bimodal in the sense that we expect something like that right, to come and not kind of something like this to come, right.

So here the means are just way too close. Not just the mean, there is also, there are also other conditions - the other thing is, one more thing that you should satisfy is that the ratio of the standard deviation the standard deviations must be small. The standard deviations must be small. So what does this mean? Okay, which actually means that, you know, they choose again some number, this is again hyper parameter less than sigma 1 by sigma 2 is less than 2. Okay, so point 5 here.

Now what this means is that, you do not want to end up in a situation where you know, you have a distribution that looks like this, I mean after you have put the classes, right, you end up with something like that, right? So you have a very small variance for this guy and then a fairly large variance for this guy. So this ratio of the standard is - so you do not want something very skewed right, towards let us say one of the classes.

So you want this ratio the standard deviations should be, to be, right, within the kind of a reasonable value. It happens to be like point 5 2. And then third, test for bimodality is that, the peak to valley rato should be greater than 1 point 2 5, again this is 1 point 2 5 is something that you can change, greater than 1 point 2 5. This is what these authors advocate in their paper. But again you could like try with like 1 point 5 or something.

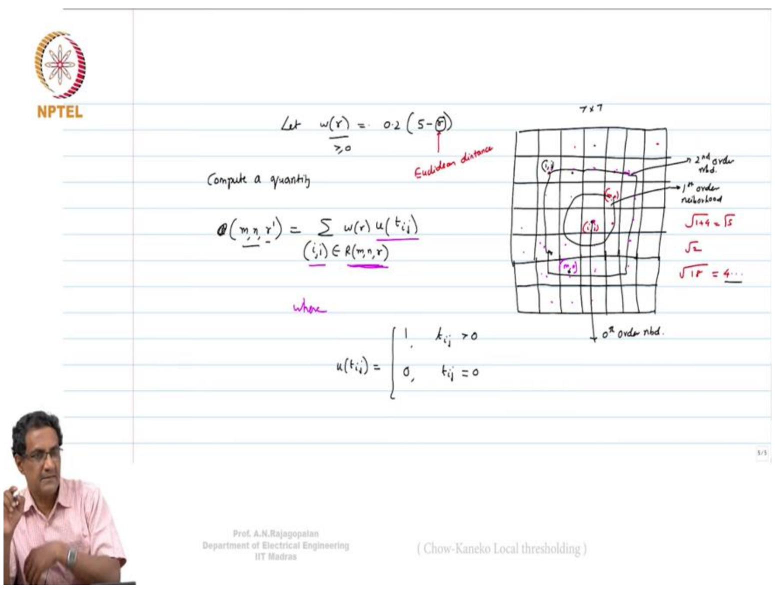
Again, up to you. What this means is that - the minimum of f of - so where, you know, where let us say, so what this means is that okay, you have the say, you have you your histogram. So f of mu 1 minimum of f of mu 1 and f of mu 2 divided by f of mu at the lowest point, okay, let us call this mu low. Okay, this should be greater than 1 point 2 5.

That means what means is that - it, assume that here is your valley okay, that is your f of mu low, so this is your valley. And then you have your peak values which are all happening here and here. This is your f of mu 1 and mu 2, right? This is the value of the histogram at mu 1 and mu 2. Now pick the lower of the two, okay, do not pick what is higher; pick the lower of the two so that is minimum of f of mu 1 so this mu 1 and this mu 2 okay, where mu 1 and m 2 are the means of these two classes - take the lower of two, divide by the valley, right, between the two which means wherever the histogram hits a low, between these two means, take that value and check whether it is greater than 1 point 2 5.

If, the idea is that if all three conditions are satisfied, if all these conditions - all, okay mind you - all these conditions are satisfied, then bimodality test is considered is passed. Okay, then it means that it would have passed the, it would have passed the bimodality test, else not. Even if one of those conditions fail, then it means that it has failed a bimodality test in which case, we will assign, for example, if this guy failed the bimodality test, we will assign t i j to be equal to, t i j to be equal to zero.

Whereas, wherever you have a bimodality test, it has passed the bimodality test, we will have t i j star there, okay. For example, here it might have passed, let us say here it is passed. So in all these places we will assign a t i j star, maybe it did not clear here and therefore t i j will be zero there, okay. So this is of course a seven cross seven kind of - these are seven cross seven regions. As I said earlier.

(Refer Slide Time: 10:38)



Now going forward, now let, now I am going to explain to you what these things mean. Let w of r the subject of a weight, be equal to point 2 into 5 minus r. Now it might actually look a little crazy to you as to why somebody would want to do something like this. The idea is as follows. Now let me just redraw that so that now I do not have to hop between slides. Now let me just draw that seven cross seven - 1, 2, 3, 4, 5, 6 and then you have 7. 1, 2, 3 , 4, 5, 6 , 7.

Similarly - 1, 2, 3, 4 then you have got 5, 6 and then... just extend this all the way down, extend this down, down, down, down down. 1, 2, 3, 4, 5, 6, 7, right. So you got like seven cross seven regions. So you divide it, the image into seven cross seven regions. Now, if you look at the center right, so the let us go down. So there is your center, okay, of this entire image. So this center, right. So when we say r equal to so here - so the r that you have here so when you say r equal to zero, that means that you are right there.

It is, and then your r could now vary, for example, for example if this is i, j and this is m, n. Okay, so when m,n is equal to i, j then r is zero. If m, n happens to be here then your r is 1. So this is the Euclidean kind of a distance, Euclidean distance. So if you are here, then it will be, it will be root of whatever right? I mean from here to here it is 1. From here to here it will be, it will be root 10, right?

So each is considered to be a unit distance. So think of this to this as - this as a unit distance and therefore, oh sorry it will be 1 plus... 1 plus 4, okay. 1 plus 4 that is equal to root 5. Okay, so something here, something here will be something like root 5. Something here will be like root 2. Then something here will have a distance as you can clearly see, right? So this has root 3 square plus 3 square - this is root 18 which is some 4 point something, right?

Okay so this 2 is that. So wherever you are, so the i, j could be anywhere, okay. So if your i, j is here, then of course your m, n would then be, then be right in and around it. So m, n could be equal to i, j. m, n could be this, could be this, could be this also. And therefore, the r will accordingly change. So it is kind of a weight. And if you notice carefully right, in their paper, they used 5 because r can go, can go up to 4 point something. It cannot cross 5, right, wherever you are in this grid.

Okay, one more thing - okay, you can only go up to a maximum of 5, so what this means is that if you have taken this to be i, j, then you are only kind of, right, allowed to go till n, r such that the r - it does not exceed 5. Okay, it should stop there. But then you could go in an asymmetric way, it is okay. For example you could be here, see if you are the center I mean, right, it could cover all the blocks. If you are somewhere, if you are not at the center, right, you may not cover all the blocks, and you may not even be symmetric.

For example, if you are here, then on this side you can go only a few whereas on the other side, you can go more. But how much you can, how far you can go - this is limited by the fact that r cannot exceed 5. Otherwise this weight will start to be a negative number. So w r is always greater than or equal to zero. Okay? And, okay, so basically that is your weight w of r. Now given that, right, this is w of r and there will be a little abuse of notation, right as I write. So I will explain what I mean.

So when you take this, so when you are at i, j and you simply consider this region, which is its own, that would be called zeroth order neighborhood, neighborhood. When you go outwards, okay, so all these guys that are surrounding this - this is all first order neighborhood. We will treat them as first order neighborhood. Then outside of that, right, all these guys that are sitting, only, only these, okay? Not, okay it would not include the guys inside. This we will call as our second order neighborhood and so on.

Okay, so it excludes, so the second order would exclude the first order and the zeroth order. The first order would exclude the zeroth order. The zeroth order includes only the region i, j. So this will be a second order neighborhood, okay. This is how we will treat this neighborhood to be. Now compute a quantity, right, what you wish to do is compute a quantity given, a quantity called theta m, n, r dash such that this is equal to summation w r, this w, r is of course, you know, the way it has been actually defined up into u of t of i, j. I will explain what u of t i,j is.

And where the summations carried over, okay, well... i, j belonging to r of m, n, I am going to say r. Okay now what this means is that, right, if you are sitting at some m, n, okay - so suppose let us say that you are sitting at some m, n here, then when you talk about r dash, so if r dash is zero then it means that, that it means that you are only kind of, right, look at you know, t m, n. So i, j - so then it will be that you are only going to be looking at t m, n. If you put r dash equal to 1, then right, you know, that would mean that you are going to kind of look at at a first order neighborhood.

So if you are here, if this is your m, n then you are going to look at t i, j, okay, let us say that this is - let me just indicate it by a different color. If this is m, n then your i, j will then involve this, this , this, this, this , this, this, this - so eight neighbors. Okay, so when you say i, j belonging to the neighborhood of m, n - so it will mean these eight neighbors. If I put r dash equal to 2 here, then it will mean that you are going to look at these fifteen neighbors. Well in this case because m, n is here so therefore, right you can only go as far as this r, right, does not exceed 5.

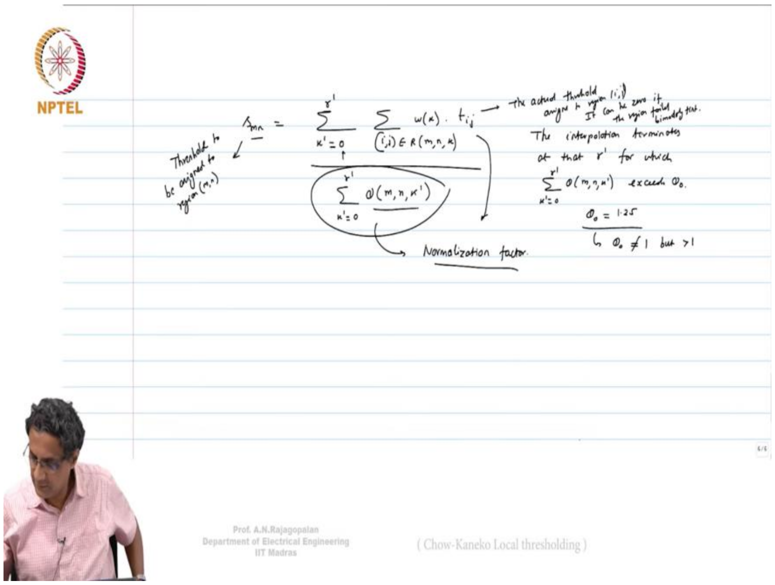
Whereas if you had taken this m, n as the center, right? If this was m, n - if this was at the center then of course it would have meant that r equal to 2 would have meant this entire fifteen, whatever sixteen blocks, right, out here. This, this, this, this, this, this, this, this, this and then all these, all these fifteen blocks around this guy. But you know, the second order neighborhood all those would be sitting here.

So the i, j - all those i, j's okay, belong to this neighborhood of m, n. Okay so whether you choose i, j or m, n, it does not really matter. But then if you choose let us say - m, n to be the region, then you will look at what are all the i, j's that should get involved right? So which i, j regions should get involved. Fine, now what is this u of t i, j now? So u of t i, j is such that, where, okay I am going to write this - u of t i, j is equal to 1 f t i, j greater than zero.

Let me write it differently. So u t i, j is equal to 1 if t i, j greater than 0. That means if a threshold had been assigned to a region i, j then u t i, j will get a value 1, else 0. That is f t i, j equal to 0. Because we know that where the bimodality test failed, we had assigned the value t i, j equal to 0. So for those things, this u t i, j will be 0.

So u of t i, j will only carry a value 0 or 1. Okay? So for example, w of r - if you put r equal to 0 you ger w of r to be equal to 1 which is the maximum weight that you can get. And then as you go outward, then this weight will start to fall. The idea being that the influence of a threshold coming from a distance which is farther off, should be less, right? So this w of r accounts for that.

(Refer Slide Time: 20:28)



Now, what you do is computing, coming to a threshold, right? So here if you are looking at some m, n. Let us say that I am here and I am looking at s of m, n that I want to assign, okay, s of m, n and please note that this threshold is assigned for all regions. Even if an threshold has been already assigned to a region, we do not simply take that. We still interpolate. Okay, so s of m, n is in fact found for all m, n. So for all 49 blocks.

S of m, n that means this is threshold to be assigned to region m, n okay - that is what it means - to be assigned to region m, n in that seven cross seven block. M, n - so this is equal to summation - this is again a slight abuse of notation. K dash equal to 0 to r dash. Then there is a double sum here. And then this is w of K into t i, j. Now t i, j okay, this is not u of t i, j - this is the actual value of t i, j which is a threshold. It could be either 0 or whatever it was assigned. If there was an optimal value that was found because it passed, this one, the bimodality test.

So i, j belonging to what do write - r of m, n, K. And divided by summation K dash equal 0 to r dash and theta of m, n, K dash. Okay this is how it looks like. Now what does this mean? What this really means is this. So if you at some m, n and you want to assign a threshold by interpolating from the neighbors. So you are going to look at the neighbors of i, j and then you are going to look first at the, at itself which is k dash equal to 0.

So the point is, if k dash and of course, you know, and one more thing I should add - the interpolation terminates, okay - here I am going to write that the interpolation terminates. This is an interpolation right? Interpolation terminates, terminates, at that r dash for which summation k dash equal to 0 to r dash theta of m, n, k dash exceeds theta knot. Now theta knot in the paper, it shows as 1 point 2 5. Now this again is a hyper parameter. You can change it.

So what, so the way, right, the way I see it is that - what they want to do is, for example, if your hand region that already had you know, a threshold, then w of, so when you put k dash equal to 0, so the immediate neighbor, the first neighbor is the zeroth order neighborhood for which w of k will be 1. Because, right, it will move all the region itself. And if then, there was a threshold, then right, you would get some value of t i, j.

Now because of the fact that now for that, theta m, n, k prime, right, as you can see from the earlier equation. If you go to the previous one, so theta is m, n, k prime. If your k prime is equal to 0, u of t i, j would have been 1 because you have already assigned the threshold. And w of r right, for the region which his inner most, when you take r dash equal to 0 is 1, therefore, right, this number is already 1.

Now if they had said theta equal to 1 point 2 5, then you know, we would have stopped the interpolation right there. That means we would have taken s m, n to be equal to, you know, to be equal to, say, t m, n itself. They do not want to do that, they want you to actually go out and scout for some more values ok t i, j which is the reason why theta knot, theta knot is not equal to 1. It is actually, but it is actually greater than 1.

So that they are forcing you to change your threshold even if you have 1. So even if for some block, you already have a threshold, right, they are not stopping there, they want you to scout for more regions in order to actually interpolate, okay. This is their idea; this is the author's idea. Now given that, right, that is the case, then if you again come back and say interpret this, so what this means is that - you keep going from k dash is equal to 0 to r dash and keep accumulating these weighted values of t i, j's.

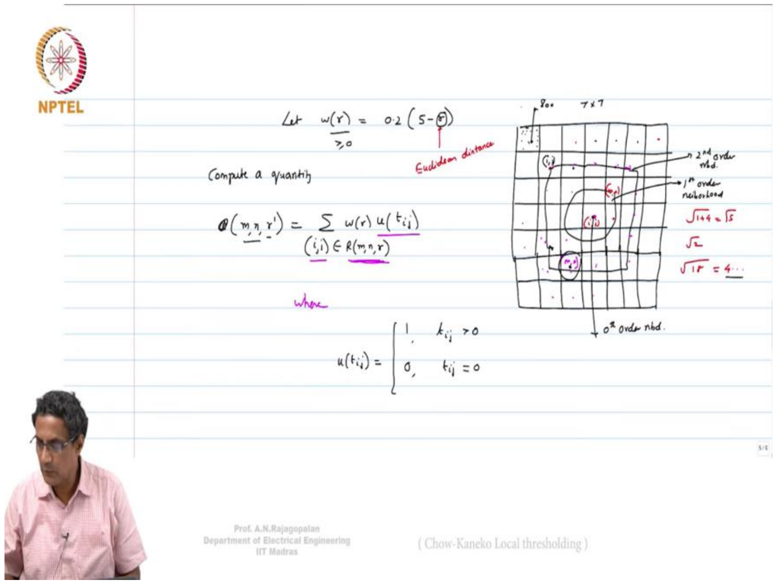
Now in some places where let us say, you have an unassigned kind of a threshold, so there you know, t i, j will be 0. Okay, so it does not mean that you will get a value t i, j everywhere. So it can also be 0 for those regions where you have not assigned any kind of threshold. This denominator, okay, so it does not mean that you compute summation till you reach r dash equal to 2 or something, 2 or 3. You will stop the moment this guy, right, exceeds 1 point 2 5. Now one more thing, right, to note is that this denominator also acts as a kind of normalization constant.

This is not only enable you to kind of terminate, you know the interpolation but it also acts as normalization factor by which what we mean is, if there are several blocks involved, right, see for example if some, if for some let us say m, n okay, where if I am sitting with my m, n there then it is very likely that I have access to several regions from where I can pick, okay. Which then means that if there is a possibility that this value will go up because I am trying to pick from several regions, but then if that happens then you do not want s m, n to go and to become unnecessarily big simply because it has access to more regions.

So you kind of balance that off by using a normalization factor so this denominator will also go up as you keep accessing more regions and therefore it will try to balance your S of m, n. On the other hand, if you have a region like this, let us say, somewhere here, let us say this is your m, n and then you do not have access to that many regions then it will mean that you can expect your numerator to be kind of a little may be a relatively smaller.

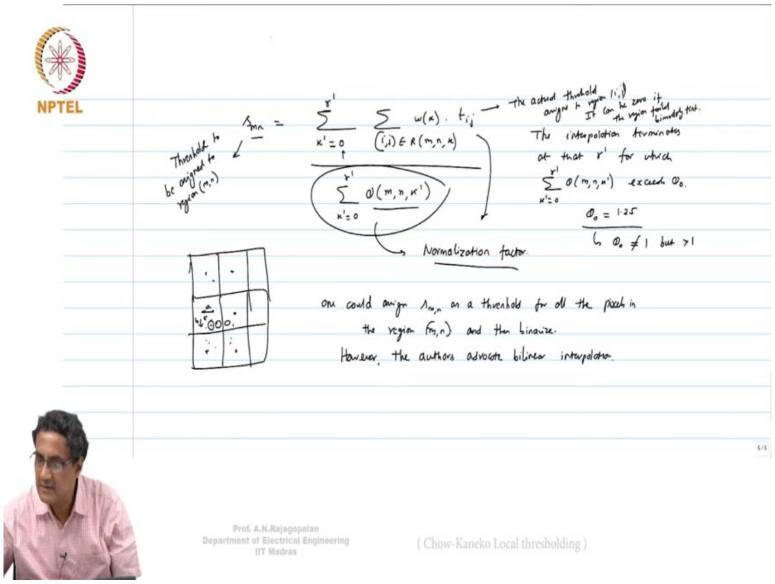
But at the same time, right, a denominator will also turn out to be relatively smaller and therefore S m, n will get kind of balanced out, right? So this denominator in this expression can also be looked upon as some kind of a normalization factor, okay. So t i, j is of course the actual threshold assigned to region t j, the actual threshold assigned to region i, j, okay? It can be zero if the region failed by modality tests. If had failed the bimodality test then you can also get actually a t i, j which is zero. And then, right, if you do this now, what you will end up is with this thing.

(Refer Slide Time: 28:16)



So now after you do this, so doing this will ensure that you get actually a threshold for each region. Every region will get a value of threshold S of m, n. Now you could assign all the pixels within this region that threshold and you could say that just use that, in the sense that then, for example, if you have assigned a threshold S of 0, 0, right, S 0, 0 for this region, then you could assign the same threshold, use the same threshold for each location here. Every this one, pixel within this region and then you could check the intensity of the pixel; if it is greater than S zero zero, we will make it 1, if is less than or equal S to zero zero, we could make it zero. Okay, this is one way. This would be simple.

(Refer Slide Time: 29:04)



But then what the authors advocate is really an interpolation. So instead of, so one could assign S m, n to as a threshold for all the pixels in the region m, n and then do a binary ration. However, the authors advocate bilinear interpolation. What does that mean? So let us say, bilinear interpolation, what we mean is since we have assigned that threshold for all the regions; suppose you are sitting right in some, you know, somewhere here and you know for this region, it could be anywhere within the grid.

Now you have values for this, you have values for this, you have values for this, you have values for this, right? For all the centers you have got an S m, n as of m, n and therefore, if you would like to, if you would like to assign a threshold for these pixels, right, the raw sitting inside this grid, then you think of this as forming a unit distance.

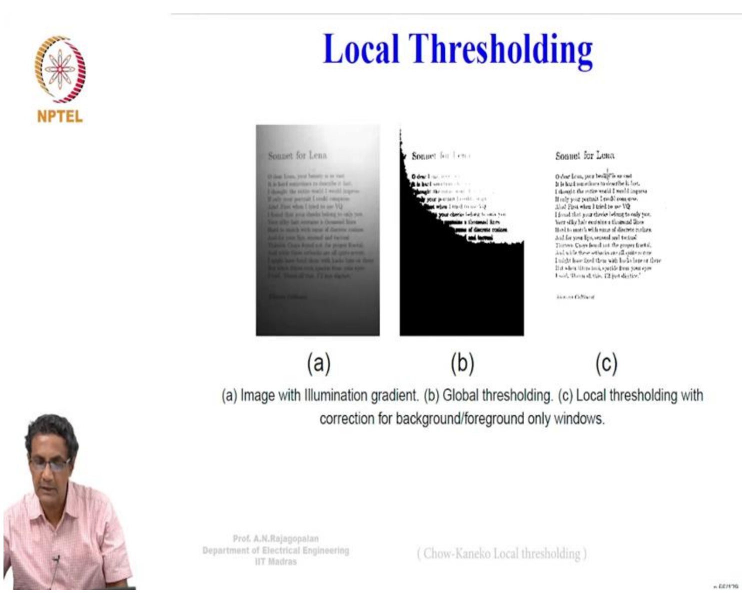
So between these two, consider this as, just as we did a bilinear interpolation, right, think of this as a unit distance, think of this as a unit distance and therefore some pixel is going to be a units away from this extreme top-most left. And then you know b down, and therefore, you know, you can actually use this pixel, this S m, n which is a threshold, this threshold, this threshold, this threshold and this threshold in order to do a bilinear interpolation.

So in the same manner that we have done intensity interpolation, so you will do like 1 minus a into 1 minus b into let us say S of m, n which is here. Then you might say plus - what would you say - the plus you would say a into 1 minus b into S m, n for this threshold, whatever is the value of S m, n here plus maybe you will say b into 1 minus a into S m, n at this location plus a into b into S m, n at this value to assign a threshold for this pixel.

And same thing, right, you can carry on and then you can kind of assign a threshold for let us say, for every pixel, right, within this seven cross seven. So for example if you go out and then if you have got blocks there, so you will have S m, n there, you already have S m, n as the center for this. Now you will look at a pixel here again take these four as your neighbors and then start to interpolate, right?

That way you can fill up t star, you can fill up the entire image with actually a threshold for every location and then do a binary ration. Okay, this is called an adaptive kind of thresholding method wherein we have extending Otsu such that, you know, the Otsu value, okay, threshold values that we have computed in different regions can then be interpolated across the entire grid.

(Refer Slide Time: 32:32)



And this example that I showed was obtained like that, okay. This one was done like that, was implemented, this was Otsu implemented with this kind of an adaptive kind of a threshold in order to implement a local scheme, a thresholding scheme which is then used to binarize the image.