Modern Computer Vision

Prof. A.N. Rajagopalan

Department of Electrical Engineering

IIT Madras

Lecture-36

Okay, so let us move on. So last class right I will do this talking about log, it stands for Laplacian of the Gaussian and we had shown that it is actually a band pass filter. We will kind of revisit this today quickly down the line, not now. And there is another thing right, so what to say, so you know there is actually an approximation that you can do of the log which is through a DOG okay and DOG stands for a difference of Gaussian. There is also another DOG which is a derivative of the Gaussian okay, but this is a difference of the Gaussian, difference of 2 Gaussians okay. It is possible and there is a theoretical explanation for this which will not go through but many times right you will find that a simpler approximation to log comes through a DOG which is a difference of Gaussians.

Of course, they have to have a sort of a relative spread, so difference of 2 Gaussians will look like some GF xy, σ-some GF xy, k σ or something okay where this k has to be chosen appropriately and such a difference also tends to approximate a log okay and for computational reasons people typically prefer a DOG. In fact, when we go through some of those feature detectors right we will see that, we will see how this DOG comes into play you know it is actually an approximation to the scale normalized log and so on. I will talk about it when we do feature detectors okay. Just a quick point regarding the Gaussian because in your assignment right the lab 2 there is some filter implementation and so on okay.

So I just wanted to get a quickly tell you how to do a Gaussian, so for example right if you think about it I mean right you might wonder that a Gaussian is sort of infinite in extent and so on, so you can do some quick approximations. So for example if you have a Gaussian which is GF m, n a discrete Gaussian especially right, so let us say you got like 1 by 2 pi σ square e raised to power - m square + n square by 2 σ square okay. So when you do a discretization, now first of all right what should be the support of this Gaussian. So what is normally done is right if you take you know a Gaussian whether it is 1D or 2D does not matter you know that of course this has an area 1 right and if you take you know 3 σ, if you go like 3 σ on either side right then you would have covered approximately 99% of the area right. So only about 1% lies outside and therefore typically what is done is you take like you know 3 σ on either side, so - 3 σ to C + 3 σ and that sort of right covers the covers more or less the extent of a Gaussian because you know you cannot have a filter that goes you know right very high this one dimension it does not make sense also.

So what is typically done is you take so suppose somebody specifies a σ for a Gaussian this is one way to approximate there are other approximation but the simplest I thought is this.

So what you can do is somebody gives you a σ then you can do a seal of 6 σ + 1 such that it is an odd number okay typically filters are odd right so that the center is clear I mean if you take a 4 by 4 filter it is not even clear where the center is right and then you do not lose anything you can just you can actually construct a Gaussian such that right when it extends you know to an odd number the size. And then what you should do is you should compute gmn for all the way from m, n right going from so going from - 3 σ to 3 σ or whatever right if this seal comes out to be let us say a 13 cross 13 right then this m and n will go from this - 6 to 6 right that will be the extent if it is 5 cross 5 then it will go from right - 2 to 2 whatever right depending upon that your m and n will take a certain range you calculate gmn from this equation but that if you sum up the gmn right that need not get a kind of sum to 1 because this is a discrete approximation therefore you simply do a some normalization right I mean that is the simplest way to normalize. So you can create a create a you know a g dash of m, n which would be equal to what if you wanted gmn to sum to 1 what would you do exactly right so simply do a some normalization of gmn and that will give you a g dash m, n which you can use an approximation for this and Gaussian okay so this is one of the things that you should be just aware of because then uniform filter is easy I mean I give you 5 cross 5 you simply say 1 by 25 everywhere but if I say you know create a Gaussian with a certain σ right then you might wonder what should be the spatial extent of the filter and so on and therefore this is one of those simplest simple approximation that you can do. And this one right and the Gaussian is not just something right that plays a role in terms of you know a weighted smoothing right so one of the nice characteristics about a Gaussian is that when you are sitting at a place right it tends to give more weightage to that pixel than to the neighbors whereas a uniform filter if you think about it it does not care about you know right it just gives equal weightage to whatever is around it but you will typically believe that let me give more weightage to where I am right to the pixel you know that I want to average and I should give more weight to that and then I should probably you know decrease my weightage as I go outside right that is what a Gaussian does so it is a kind of a weighted averaging right as opposed to a uniform filter which is like you know uniform weights all over and this control that you have right the σ gives you a nice control and this as there is a whole theory called the scale space theory okay that is based upon a Gaussian okay we will talk about some of that okay.

So there is something called the scale space theory where the scale is associated with σ right and then you start examining an image at various scales and then the idea is that you know certain things pop up you know depending upon how you kind of right how you run through your σ and so on that is something that we will see later but just wanted to tell you that Gaussian does not end of the story for the Gaussian. And as we go forward you will keep encountering Gaussian in you know again and again. Then of course you know I talked about horizontal and vertical edge filters right something very simple that we talked about last time and the smoothing filter okay is a Gaussian so it is kind of a weighted smoothing and let me just show you some examples right just to kind of give you an idea and then we will go to the okay this I will come to later okay. So say a box filter when you smooth right this is how you will get because this is like you know smoothing everything right equally and this is again right another kind of a box filter result again right another whereas if you use a Gaussian right then those are all skip I already talked about all this then you see that right it is not exactly the same as what you get with an averaging filter. Of course it does not mean that we do all right everything with a Gaussian and all I mean these days there are these are all called local means okay in the sense that you are averaging locally okay now there are filters that are called non-local means that means you do not have to necessarily look around you can actually go scout for values elsewhere we will not be talking about those kind of filters right in this course.

Then this is a uniform filter again and then right you can see the effect of effect of σ here I think so it is like this so as you go down right you are kind of you are kind of increasing the smoothing so around increasing the smoothness so if you take any one okay if I tap here you can see that if you take any one and then if you go down right it is like you know no smoothing then σ equal to 1 so that will be like what will be the spatial extent σ equal to 1 7 cross 7 right 6 σ + 1 right so that will be like a 7 cross 7 filter which you will build up σ equal to 2 what will be the spatial extent okay. So you can see that you know as you go down right I mean you get the blurring effect to be higher and higher and then you take a Gaussian filtering then depending upon what value of σ you choose right you can end up blurring things heavily and sometimes this could be used for let us say privacy and so on but not always okay there are various reasons why you want to do smoothing and then a Gaussian versus box filtering right this hopefully brings out some of those advantages that let us say right a Gaussian has so you can see that you know the bricks are I think the bricks are actually right I mean if you look at the top one top right that has a better sort of filtering right in the sense that you can see that the central weighting is actually better there compared to something like a you know uniform filter down. In this is a gradient kind of thing so what do you think is the second one so that is the original image is the first one and what is coming in the middle so what kind of a filter is that what kind of edges are you seeing there vertical right so therefore it is a kind of you know a vertical filter and then the other one is actually trying to pick up all the horizontal edges so it is like a horizontal filter we will also talk about orientation of edges and so on down the line this is again another example for Sobel so again that you have a brick and then depending upon which filter you apply you can pick up horizontal or kind of vertical edges and these edges are something that we saw in CNNs also right okay this let us kind of skip this for the time being this and all we will skip and then yeah so well I will kind of come to this when we do you know edge detection a derivative of Gaussian this again I will come to later okay yeah I think at this point of time right let me go back to whatever is talking and okay so the other thing right that I wanted sorry yeah. So you have written e power - m square +. Well this is all courtesy zero centered the one that I have put is actually centered about the zero and normally right that is the way you look at it I mean you have a Gaussian know I mean so you take the central thing to be kind of say 0, 0 and then you can extend center is always at 0, 0 the center is always 0, 0 even in a uniform filter if you have - 3 to I mean you will always take it to be - 3 to 3 and the center is always kind of say treated as 0.

So it is like you know it is like in that filtering operation right if you look at it you have like gmn and then let us say is equal to some you know summation fmn h of m - n or whatever so that h of m - n or you write it as hmn f of so that hmn is always zero centered. So it is like you go from the you go over the spatial extent of the filter and that is over that is about 0, 0 always yeah so here we do not have that you are probably thinking in terms of mean and things like that right I mean zero mean I mean if you want to think about but it is not statistically in that sense this is like a deterministic filter right. And there is just this one more thing right I see based upon what you have seen right there is something called unsharp masking okay which is what do you think it does right because always right people get this is a very strange term unsharp masking what do you think it must be doing unsharp masking by just look up this word right what do you think it must be doing is it like sharpening or is it like smoothing right no it is actually sharpening it is like unsharp that means you are smoothing masking that means you are masking the blur right you have to think about it like that I do not know who coined this word they would have made it very simple but that is what it is called okay it is a very standard thing and the way this works is as follows right so you have i suppose you have an image i right then the unsharp masking looks like this so the original i + a constant times whatever a or k or whatever i - i blurred and this guy is something like 0.2 and between 0.

2 and 0.7 again you know there is nothing sacred about that. So it is like this right so you take so again Gaussian comes into play here or any kind of smoothness filter so you blur i to get your ib and then you do i - ib that is like a detail that you are getting right that is like a high frequency information and that you kind of you scale it right by some number and then the scaled high frequency detail you add to the original image right that is why it is called like unsharp masking so it is actually sharpening effectively I mean you are trying to bring out a detail and that is what actually was there in the first one right when you if you saw you noticed the very first one right that the very first thing right so what does blurring so if you see right this is the original Lena then there is a smooth Lena here it is just a weighted average this uniform filter then this is a detail so if you subtract that i - ib that is the detail and then you multiply it with some constant this detail and then you add it to the original right then you end up actually sharpening sharpening the image so this operation is called unsharp masking. i - ib blur will be a sharpened version. No it is not a sharpened version of the image that will be a that will just contain details high frequency detail edges and so on that would not be sharpened image itself that will be like I think and then you add it to the original. Sir in washing filter it seems like that they are taking some features but they are not so.

Yeah it is like this right see when you do a blurring right and so it is like this if I gave you let us say right a low pass filter and if I said that I wanted a high pass filter out of it how would you arrive at that suppose I had a filter already which can do a low pass filtering operation and I want a high pass filter right what would that be. No no no I am saying just a filter okay let us not even worry about an image let us have a filter this can do a low pass filtering operation but I want from this filter I want to arrive at a high pass filter. Be very careful it is in the spatial domain right okay right in time domain for example if you wish or a filter in the spatial domain 2D filter I have this if I run around I can smooth correct now I want a high pass filter I have this I should be able to get a high pass filter from this right how would I do that. No but then in spatial domain I am asking the spatial domain right what would that be all pass yeah what will that be what would it do a delta mn - this would that not be would that not be the one right you see in the Fourier domain if you think about it you need something like a 1 all through - the low pass that will give you the high pass I mean right like this no I mean I mean you have a low pass I want a high pass right so what will I do I can do this 1 - this that will give me a high pass what is what is the 1 for in the spatial domain it is a delta right. So in the spatial domain we will do delta which will be like just a delta at the center everywhere it is 0 - this right that is what you will do that is exactly what is happening here no I - Ib if you look at it Ib is what Ib is simply I convolved with some low pass filter so you can pull the I out you will have well I mean right in this case because it is actually a convolution right it will be a delta - - the low pass filter that is what it is so it is not like you are losing anything right I mean you are getting high frequency information details.

So unsharp marking right so it works like that and now right this actually the other application right of these filters is an actually is actually in you know this one denoising like I said again that simple applications these are not very complicated I mean if you look at the reason most recent literature and all it is way beyond what we are talking about but I think it is important to at least know these things so that when you read that you know you can associate you can relate to what we have done. So denoising the simplest case right that you can think about is actually AWGN so what you have is an image right with you know which has been affected by let us say Gaussian noise okay and let us say with a kind of you know with the same variance everywhere and you want to be able to reduce the effect of noise right. Now there is I mean you would have all seen that you know people typically will take a filter and then right and then you know you do a spatial averaging but that comes out of actually you know I will just run through this very fast okay so it is like this I mean if I gave you multiple frames right this actually happens in your cameras see if I gave you multiple frames of the same scene quickly taken right so that you know there is no motion nothing happens I just grab you know 50 frames quickly can I sort of do something to you know mitigate the noise what would you do? You will take the average of all of them right because the scene has not changed I have got multiple shots of the same scene very quickly taken so you would do an average of that right and what will be the effective variance then of the noise you see originally let us say each one has a σ square noise right in it I have taken L number of images I have averaged them what will be the new variance? Σ square by L whatever right so this is the right number of frames so that comes out of this right so you have like gmn and this is actually done in your cameras okay so gmn is let us say okay let me take the ith frame I of f of m, n + let us say eta i which is the noise of m, n right this is what you have as a model now what you are saying is you are trying to get an estimate of at any m, n right this is valid at any m, n you will just say right I will just do an average of 1 by L let us say summation gi of m, n i running from 1 to L this is what you would this is what you said right you will do. Now right this being noisy right so it means like you know you are having an estimator of f right that is what you are this f hat is like an estimator of f so one of the things that you look for is what is the mean value of this is f hat what should it ideally be f itself right fmn and that you can easily show because this is 0 mean right so this we have taken to be 0 mean with some σ square let us say noise and therefore right if you just run it through the summation above you will get that you know this is equal to fmn which is good such an estimator is called unbiased right this is unbiased. And then what you can do is if you look at the variance right how would you like the variance to be should be low right I mean about this value because you are already like the on the average you are getting to the correct value and the variance about this should be as less as possible right and therefore if you look at variance of f hat of m, n right that will be whatever right so that will be expectation what is it f hat m, n - this guy right what is that the mu f right which is actually fmn the mean right square and this if you do right it is very straight forward to show this you just substitute for f hat of m, f hat of m, n which is 1 by L this one gi and then if you do this right you end up with exactly what you said σ square n by L okay that is what you will get to be the effective variance and such an estimator is called a consistent estimator okay.

So that is the most nice thing that you can get as an estimator something that is unbiased and something you know which is also you know consistent in the sense that the more observations you add the effective variance of the noise goes down but then in reality right you cannot denoise like this right I mean suppose normally you will have just one image you may not have you know 20 images with you therefore what you do is you do what is called a local averaging and then and then the essence of that comes from here it is like saying that see here one of the assumptions which you made was that was that you know these were all aligned right in the sense. So when you take just one image and then you want to do some kind of noise mitigation then what you do is you know you take a filter and then you apply it locally and this interpretation is that the other pixels are like the other frames you know the information coming from the other frames it is like but then you have no other go now you can only locally average you do not have information from other frames you have only one frame but this averaging notion actually comes from there and the idea is that if and then the idea is that locally right these intensities will be like you know similar and therefore they are all affected by let us say independent noise right each one and therefore when you average them you should effectively get a variance that is actually less than the variance affecting a single pixel right. So the idea exactly goes like that but in reality right it may not work out that well because these intensities need not always be the same they are right you cannot assume that they are the same and therefore it affects your you know edge preserving capability because it could be that right there lies an edge right and then when you sum up right all this assumption that you are making is all gone then therefore when you do that you will get a get a smearing effect which is not which is kind of annoying when you see that because such things happen that is why you go for you go for filters that are far more superior than this okay but you know this was what was done right originally and then and then you know and then there are there are several other you know improved versions of that okay. So which is which is which is what which is which is what you can do if you had an if you had an AWGN kind of noise then the other kind of noise that I am talking about this because in that in that in your lab too right you have these things okay as part of the assignment so you have to actually you have to actually incorporate these these filters and so on. The other one right that you can do is a salt pepper noise okay so instead of instead of the right instead of the AWGN the other one like I said that something that is sprinkled and then you know which is which is kind of signal independent typically and there it what you do what you use is actually a median filter okay you know median filter is very simple so what you do is suppose you choose a size for the filter let us say I choose something like 3 cross 3 then what you do is in the image right wherever you are okay wherever you are if you take if you take a 3 cross 3 window then you actually take all those intensity values put them in ascending order or you can put them in whatever a descending order and then and then you simply pick the median median value and then you replace the central guy with that.

See in all of this you will put put the put every value in a kind of a new grid okay you do not put it back there okay you understand it when you do a convolution you always copy it into a new array I mean you do not put it back in the same place otherwise it will affect the next result right because it is all spatially related no okay so you so you need to so that and all I am assuming you know right so you kind of copy it into a new array and wherever you will have and then then in a median filter the assumption is that if there is an outlier right it will go either to this end or to that end right it will go either to the max or the min and therefore when you take the median right it should not be there right and there are various higher versions of this you can also do a weighted median just like you have a Gaussian which is doing a weighting right you also have what is called a weighted median where what you do is if you have a if you have a median filter let us say right I mean you know you just have to specify the size because in a median there is nothing like weighting or something but that weight you can bring it by saying the by kind of specifying another kernel where you say that where you where you kind of write you know put some numbers okay what these numbers mean is that each of these should should repeat that many times when you do the right before you take the median right so it is like you are emphasizing certain guys more so you say that because you know if you take if you were to take a simple median you would just you just have that value appearing once but suppose I write 2 that means you have to put that value twice and then if I say 3 that means the next value whatever is sitting here should be copied thrice and then you sort right so when you do that sorting right depending upon which one is repeating more it will have an influence on the final result so that is also called a weighted median. Now all these are all these are typically see the other one that we talked about which is a which is the Gaussian right there I said that or whatever write a box filter it was easy to come up with a high pass filter right because it is all a convolution operation whereas median is actually a non-linear filter okay this operation is you all know right median itself is not really a linear operation. So therefore right if I asked you how would you do a high pass filter with a with a median the low pass thing it is clear fairly like that if I wanted to pull out noise or mitigate noise I can do something like this and walk away but I but if I asked you right can I do a high pass filter in the sense that you know one simple explanation could be that if I had uniform values everywhere okay this we keep talking about right you have a wall or something where let us say intensities are all the same and if I apply a median filter there of course you know it will if all the intensities are the same you will again get back the same value itself right. But if I said that the output should be 0 because now it is a kind of a plane wall right it is a homogeneous region I need high pass filter what would a high pass filter have done if you are doing a convolution and all that would have simply picked up a 0 right we saw know the sum of the weights for a for a gradient filter is 0 or a Laplacian for that matter is 0. So anytime you apply it on a place that is like you know all same values it will always go to 0 how will I make a make a median filter output is output act like a high pass filter? No, no I want I want I want directly a filter well you are saying I will I will first you know create a low pass filter version but that is again right that is all assuming that there is a convolutional model underneath what you are saying right is still banking on some kind of a convolutional model.

I mean it is like I-Ib you know right so if you if you rewrite it you are coming back to convolution kind of a model but actually this is a nonlinear filter. I mean implicit I mean you are not explicit but implicitly you are implying a convolution kind of operation. How do you do that? Anyway right I am going to I am going to leave leave this to you as an exercise right think about it I might I might ask it in the exam also I want you guys to think about it. So that let us let us let us let us leave you guys to have some fun. Okay now this okay now you know all this was in the spatial domain right and exactly.