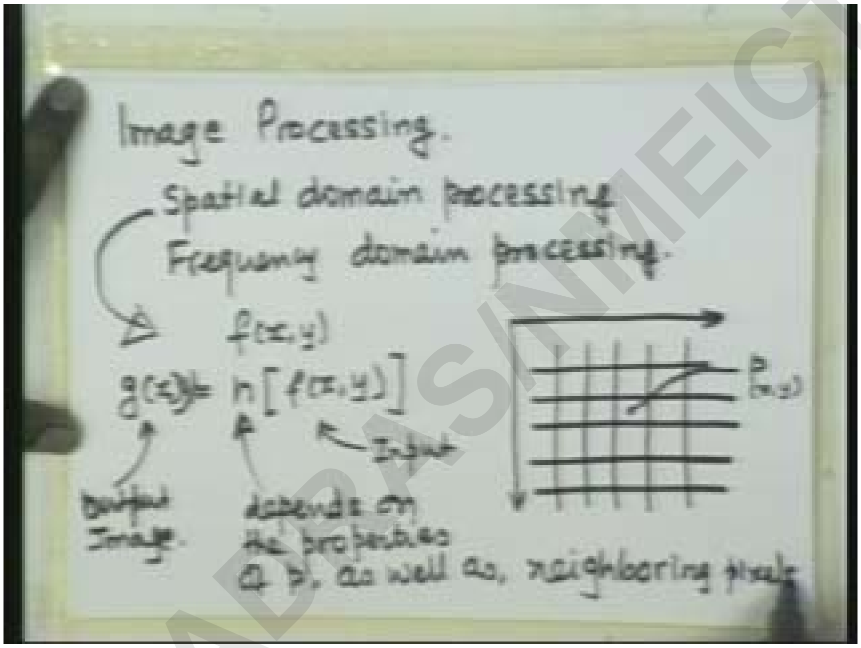
ROBOTICS
Prof. B.Seth
Department of Mechanical Engineering
IIT Bombay
Lecture No-28
Image Processing

Last lecture we look at image processing issues so um image processing basically there are two approaches(refer slide time04:06)



one uses for image processing one is called spatial one is spatial domain processing and we have the frequency domain processing

okay now before going into details of either of these first I would like to just give you a flavour of what each of these consists of so spatial domain processing basically deals with um neighborhood properties of the pixels so if I have a image so Image is basically a matrix of intensities

okay if I take one particular pixel here say pixel p with some coordinates x and y then basically my image in the spatial domain is represented by this function here for all different values of x and y belonging to the image

now the processing basically takes this image and applies a function or filter to it to yield a different image

so if I the processed image or pre processed image is sorry x and y both g x y then it is some function of the original image f x y processed by transform or a filter you can call in general this being our input image and this is the output image this function h basically depends on depends on the properties of the neighboring pixels properties of p as well as the neighboring pixels so typically what is done is that you take you have now one pixel of interest for now and we have our image which is other neighboring pixels around it

so supposing I take a three by three neighborhood of this then I basically have pixels left right top bottom of it and I have eight neighboring pixels for the pixel of interest p okay and this if I assign it some weightages so if I give this weightage w one w two w three etcetera etcetera I have w four w seven w eight w nine okay

so I have assigned some weights to each of these cells this I will call a template ofcourse other values are also there if its I hope its clear okay these are basically weights wone wtwo wthree wfour wfive wsix and so on so

so one of the preprocessing steps in general we can show this as that are transformed image or processed image essentially is obtained by summing over all the neighborhood pixels of interest okay

so if I say sum w j k times the intensity at x plus j y plus k

okay this is to be summed over all the js lets say in this particular case going from minus one zero one and k going from minus one zero one

okay if I spell it out basically this means that I am saying weight minus one minus one then I have to take the intensity x minus one y minus one right then I will take the next summation next term which will be minus one lets say and zero which will have to be multiplied by x minus one y

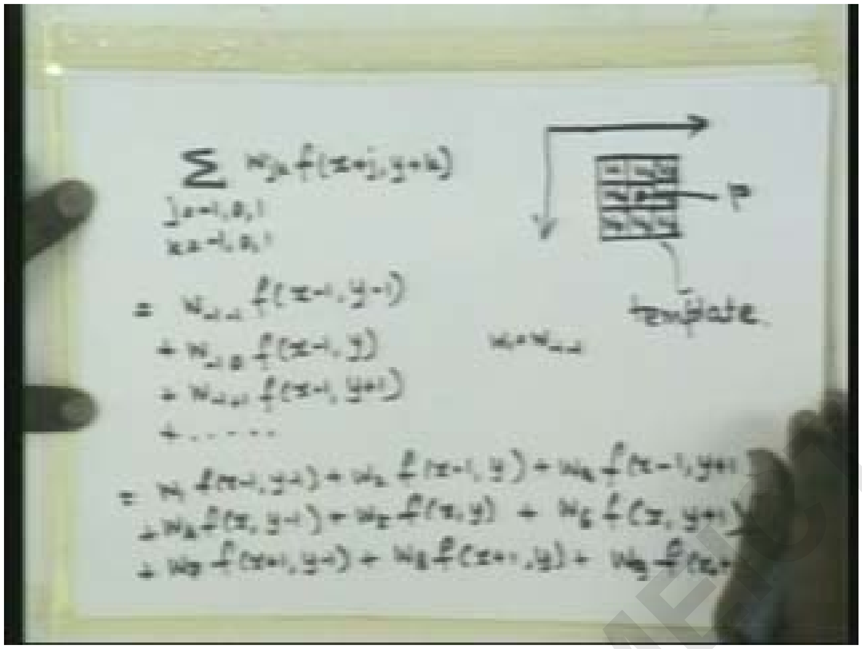
so likewise I can go plus one x minus one y plus one okay and so and so I will have nine terms like this where what I have shown here is this w one

for example is w minus one minus one because it is corresponding to x minus one with relation to the pixel that we are looking at so the idea is that you look at this particular mask or template you take it from each take it on each of the pixels of the image starting from lets say upper left to the bottom right sequentially you place this mask placing the mask has the meaning of that you multiply the way to the intensity and then find out the sum and that is going to be the new value that you should assign to the value to the pixel p

okay the way you will do this is that you will go ahead and do perform this operation throughout the whole image and then after that you will replace all the new values

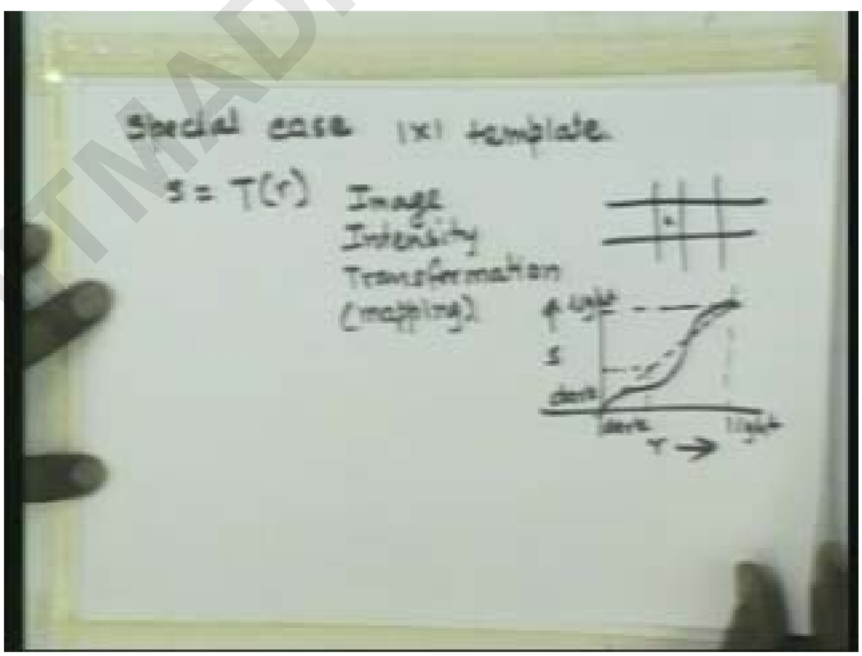
so the values will not be replaced as soon as they are calculated but they will be replaced only when the entire image has been gone over by the template

so this one for example is going to be simply w one times f x minus(refer slide time08:56)



one y minus one plus w two times f x minus one y plus w two times w three times minus one y plus one w four I am writing in the same fashion as the x y minus one plus w five f x y plus w six f x y plus one plus w seven f x plus one y minus one plus w eight f x plus one y plus w nine f x plus one y plus one

okay this is the meaning of a template this is also called a filter so depending on what ways are assigned these filters have different effect on the image we will see some of these filters as we go along now one special case is worth mentioning special case is(refer slide time11:39)



when the window that I am looking at the template that I am looking at is also one by one so one by one template which means that for the image I am looking at the intensity of this particular pixel and I am going to it does not depend on any neighboring pixels it only depends on this particular intensity value

so I can write this as some transformation of whatever this intensity r is gets transformed into a different intensity which I am calling s okay

so this is called image intensity transformation transformation or mapping okay this is a special case let us see supposing what I am talking about is you have some kind of relationship between the input intensities

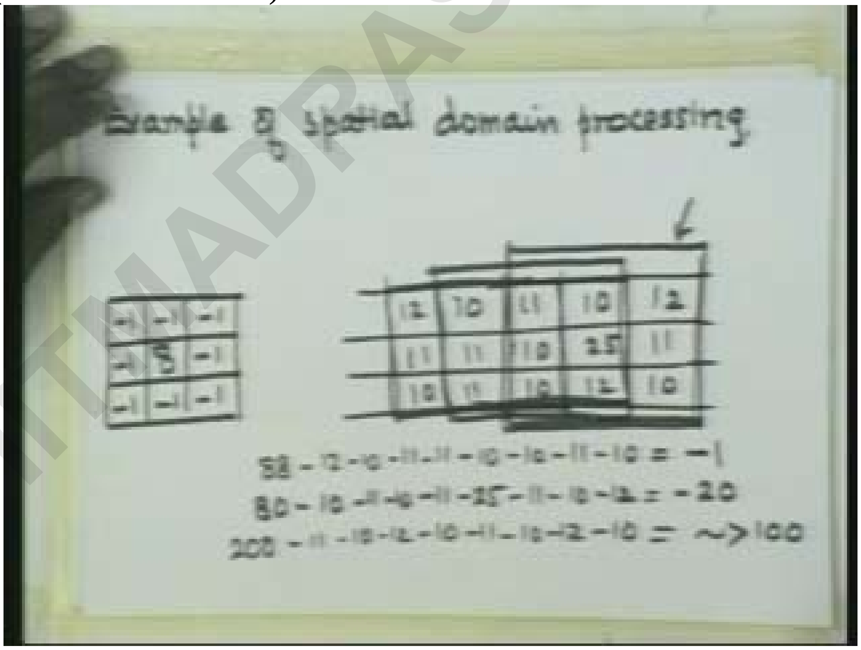
so I am plotting r here and s here so the value of s depends in some fashion to the input intensity right so this is lets say dark dark area this is the light area here again I have dark and light

so the value of the intensity is zero here lets say two hundred fifty five corresponding to the fully bright so what is happening is um supposing now I look at if I had a straight line here then the straight line would correspond to no change in the image at all right because whatever the intensity corresponding to that I will get the same value for the output image

so I would have no done nothing at all but if it is deviating from the straight line then it is going to transform the image in some fashion and this can be used for increasing the contrast or decreasing the contrast of the image okay

so we will look at this also this is also part of the spatial space space techniques

now let us apply one of these masks as example so this is example of spatial domain processing(refer slide time15:20)



okay let me take a mask define a mask first or template okay

lets say I have put minus one minus one minus one minus one eight minus one minus one minus one minus one supposing this is my template and then I take my image part of the image okay so let me put some values say this is twelve intensity ten eleven ten twelve lets say this is eleven eleven ten twenty five eleven

okay ten eleven ten twelve ten I am just this is typically what the image will be in the computer right this is the set of intensities which are there in different pixel locations so now if I fit this mask on the three by three area here okay

so what I will be doing is essentially multiplying the middle pixel intensity by eight so I will have eighty eight and from that I will have to subtract out all these things so twelve minus ten minus eleven minus eleven minus ten minus ten minus eleven minus ten

so whatever that comes out to be may be one of it we can calculate and we can find out what this is

now if I move this mask further I will translate it to the next set of three by three matrix here this should come out to be pretty close to zero minus one okay

alright now you look at this area now I am going to make it little darker here this is where I have placed the template and I am going to now apply the mask to this

so the next value for this particular cell I am getting is eight zero minus ten minus eleven minus ten minus eleven minus twenty five minus eleven minus ten minus twelve

so this will turn out to be some value may be around minus twenty or something minus twenty

okay when when I shift it further so this is the area now I am looking at and so I will have twenty five in the center here which gets multiplied by eight

so I have two hundred minus eleven minus ten minus twelve minus ten minus eleven minus ten minus twelve minus ten and this will turn out to be more than hundred

okay so what I am trying to say is that depending on the mask so this is the mask that we chosen in which the weightage is very high in the centre okay

now that has a tendency to highlight anything that differs from the background

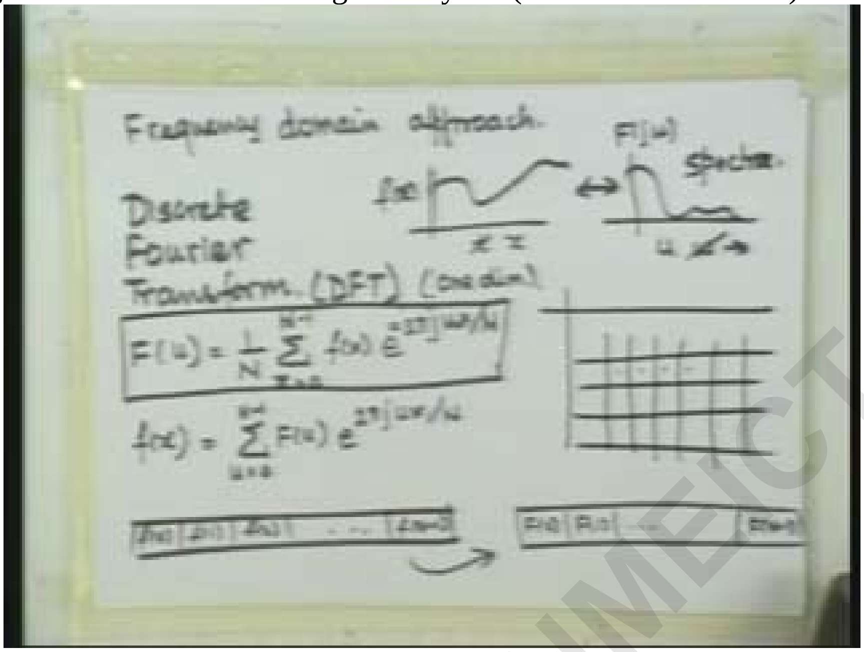
so if a single pixel is has much different intensity then this mask will highlight that because if the intensity was pretty much uniform then you will come out to be close to zero you know within some band of zero you will come out but if the intensity is one particular pack is different which happens to then lie on the centre of the mass then we are going to have a very great value a very large value so it can be used to detect lets say isolated points which could be due to some extraneous effects say could be some noise

so one black pixel has suddenly become has white for some reason and this will highlight that and if one can detect it like this one can also correct it so thats the whole idea of preprocessing

okay lets get an idea about the frequency domain approach yes please thats right so what you do is you pad it with some numbers so it could be you duplicate the first row as a minus or zeroth row or you can fill it with zeroes you know anyway edges are not where you are going to do much of the processing you will do the processing but you not expecting the features to be there at the very edge if it is the case then you will not detect it very well

no no the television is not doing this kind of processing atleast the older television did not do any such processing this you know showed the signal directory so only when you are trying to image process and try to enhance something or remove noise automatically then then you have these things likewise

okay so we are now going to understand appreciate what frequency domain okay methods or approach is okay frequency domain approach basically uses fourier transformations okay as you know that if you are given some time domain signal okay this(refer slide time22:00)



is say some f of t as a function of t it can be represented as a equivalent or alternate way of representing some kind of frequency okay

so this is now what is called f of j omega versus omega where omega is has the units of frequency which is one over time and this is some kind of amplitude and phase of this is a complex quantity

so it will have a magnitude and a phase in general so depending on the shape of this you will get different frequency content or spectra so this is also called spectra

now it represents the same amount of information so now we can do some processing which is in this domain in the frequency domain rather than in the time domain

only thing is now we have a image which is got a series of discrete numbers in each of these cells which is representing the intensity

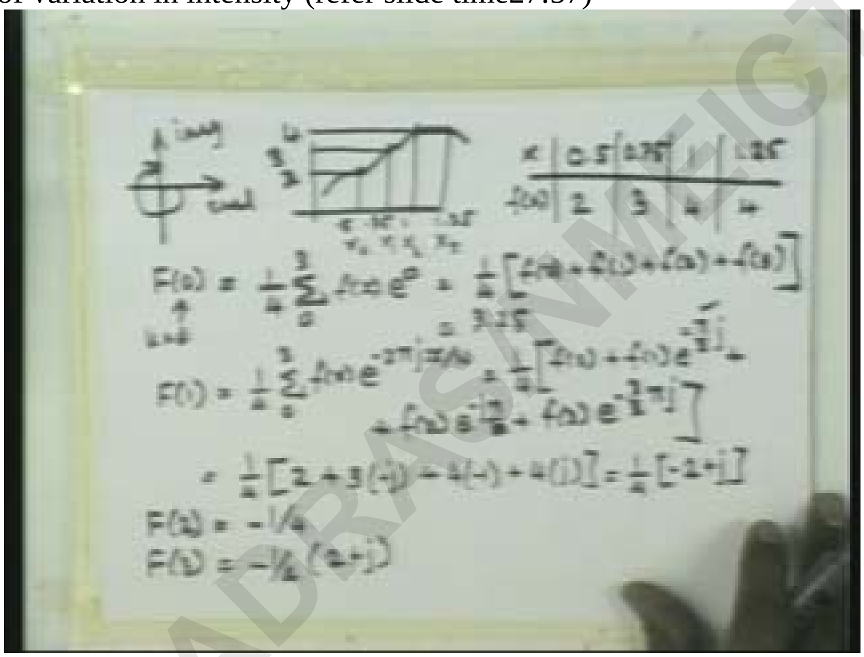
so instead of looking at the time we are going to look at space how does it in the space how is it varying instead of time changes we are looking at changes with respect to increase in one of the variables like x or y right so and it is also discrete so we have to define a discrete fourier transform

okay this is also called d f t now d f t so I can define a function which is a function of frequency u as one over n times summation of x going from zero to n minus one f of x and now defining a one dimensional first times exponential with imaginary value so j is square root of minus one times u of x u times x divided by n

okay so if you look at this this is what is known as discrete fourier transform of a function of x right and when the transform is applied on this particular function you get a function new function which is the function of frequency

so basically this is what we have done it started out with now in place of t we have x and this is f of x and now equivalent to that we are finding the frequency domain representation in terms of a spectrum which will be here we have u instead of w and this is a f of u okay

similarly we can do the inverse function or inverse transformation which can be written as u going from zero to and minus one capital f u which is the spectral representation e to the power negative of what we have put here so that is going to be two pie j u x divided by n okay
so what I have is basically if I have a one dimensional image so one dimensional image is basically pixels arranged in a one dimensional array each having some number f zero f one f two etcetera etcetera f n minus one when I apply this transformation then I will get a new set of numbers now this is going to have f of zero f of one like that upto f of n minus one where now these arguments here are representing some kind of frequency variation right
so I think the best thing to do is to look at an example and so let us look at a example of some kind of variation in intensity (refer slide time27:57)



and some value here some value here some value here and some value here
so let me say this is my x not x one x two x three which may be some intensity values lets say this is “point five” this is “point seven five” this is one this is “one point two five” whatever intensity is right and I have to know what these values are so lets say these are some two this is three this is four simplifying taking case where we have integeral values okay so now what I have is basically I have just to write it more clearly x versus f of x value of x is “zero point five” then I have intensity is two “zero point seven five” intensity is three one intensity is four “one point two five” intensity is four it’s a very small image
so that we can work it out so if I want to find out my first zero frequency magnitude then I find that this will be one over n if you look at the formula summation going from n is equal to zero or x is equal to zero to three okay
f of x e to the power what do I have for u u is equal to zero right so exponential has two pie j minus two pie j x times u divided by n so this becomes all zero for all the terms and therefore I get one by four f zero plus f one plus f two plus f three

okay that is coming out to be “three point two five” now we take q is equal to one then we will get f of one and that is one by four summation zero to three f x e to the power now I have minus two pie j

okay times one which is u times x divided by n where x is varying from zero to three so here what I get is one by four times f of zero now when I take f of zero this term will exponential will become zero so we get one so this is all we have then we have f of one times now something which is not zero and that is e to the power minus pie by two times j

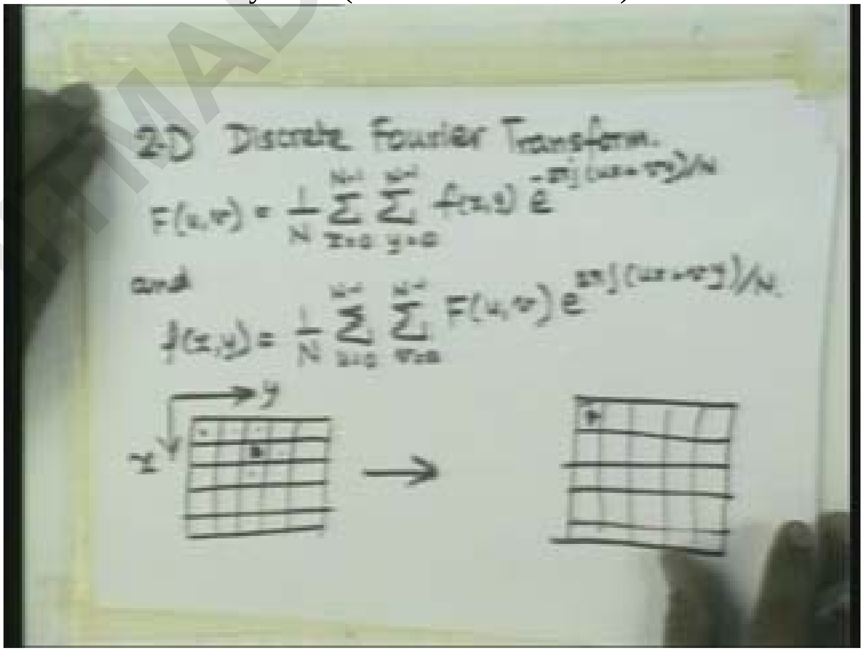
okay likewise next term is going to be so I am continuing here f two times e to the power minus j pie plus f three times e to the power minus three by two pie j okay

so I substitute the values what I get is one by four two which is the first that f zero is two plus three times this quantity which happens to be nothing but minus j right e to the power of minus j I would you know that twelve you should quickly see that this is my real axis and this is the imaginary axis and what we are talking about is e to the power of minus five sorry pi by two yeah that means correct that oh that is pi here we are talking about this term okay

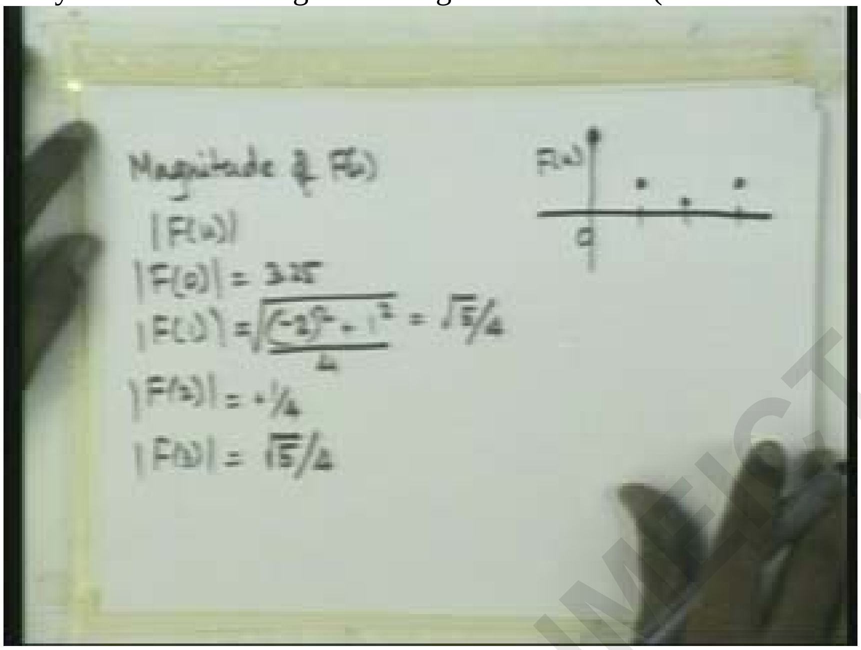
so it is minus pi by two so minus pi by two is here which is nothing but minus j right so that is there plus four times now e to the power of minus j pi is nothing but one or minus one so minus one plus four times

okay three pi by two minus is coming upto here so that is positive j so that is times j so this gives me one over four this is minus four this is two so this I have minus two and this is minus three j this is plus four j so I have j so this is the quantity that I have okay likewise without actually labeling over it let us say that we find f two which turns out to be minus one over four and f three turns out to be calculates out to be minus one over four times two plus j okay

so now I have determined all the coefficients of or all the values of the spectral spectrum which is represented by capital f zero capital f one capital f two and capital f three okay so I can plot that now this is my f of u(refer slide time33:28)



versus u is equal to zero one two three I have but I will have to plot two items so what is usually done is you look at the magnitude magnitude of f of u (refer slide time 30:03)

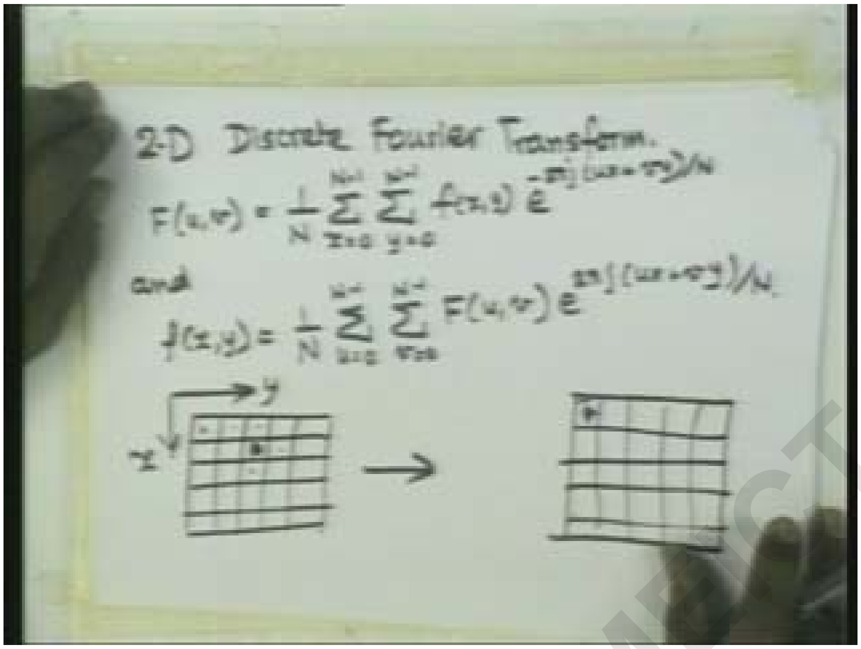


so we are talking about this quantity and so for f zero this is nothing but “three point two five” okay

only real value was there f of one is going to be now because there are real and imaginary terms so minus two squared plus one squared square root okay

so that is square root of five that is the magnitude by four is there yes and likewise I have f two which is nothing but minus one by four and the magnitude is going to be plus okay because we are taking the absolute value and f three is going to have a magnitude of same again square root of five over four so I plot it here I will have a point here then I have a point which is slightly more than one so somewhere down here then quarter and then we have same value here so this is the spectral representation of that particular image okay ofcourse one can go from this back to the original set up intensity by using the inverse discrete fourier transform okay

now this can be generalized for a two dimensional case so so we have two d discrete fourier transform okay (refer slide time33:28)



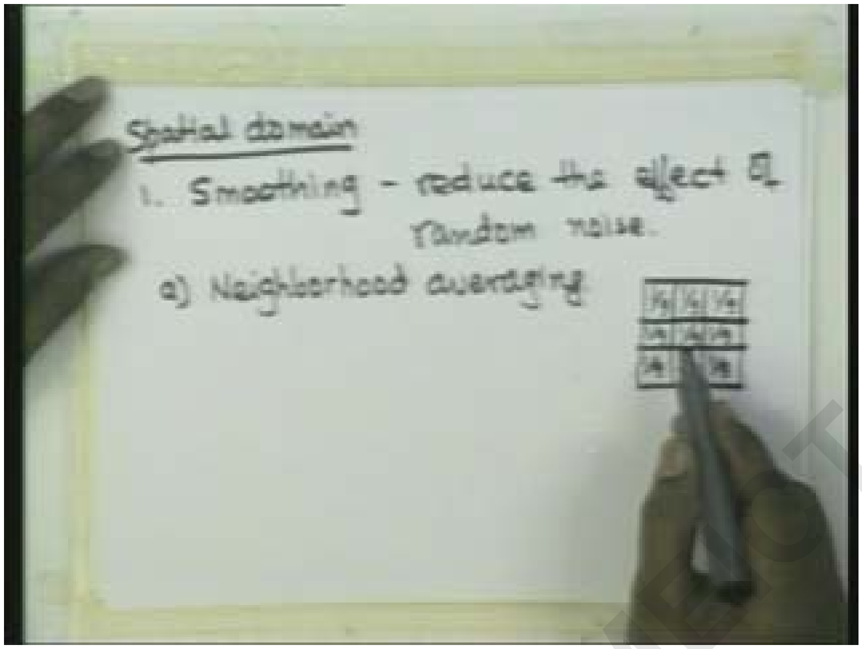
and in this case basically let me just write it down we will not do any example because that will take too much time now we have two frequency variables u and v that is equal to one over n summation x going from zero to n minus one summation y going from zero to n minus one f of x y which is the intensity at that particular location e to the power of minus two pi j u x plus v y divided by n and

the inverse will be give your intensity back from the fourier transform values u going from zero to n minus one actually you can define it for not n by n image but it could be some n by m or rectangular image instead of a square image that is also possible but lets not worry too much about it concept wise we are able to get this and this is going to be u v now the positive exponent two pie j u x plus v y divide by n okay

so here we are talking about now a two dimensional image of intensities so we have some some values of intensities at each of these cells and this transform will take me so this is my y and this is my x the way I normally keep then I will get set of intensities here or set of courier magnitudes lets say it will represent the strength of the fourier term and I can represent that as a intensity and I can plot it

so this can also be shown the transform when I do the transform my original image will now look like a totally different image of patterns okay

and if you are used to seeing it you will be able to say yes this feature belongs to that the key point here is that each of the value that I determine here for any of the cells depends on all the values of the original image right this cannot be computed from this alone it depends on the each of the intensity value of the entire image in different weightages which depends on the exponent that we have which determines this value right and similarly the other way round also you can say that if you want to transform back and find out what the intensity here should be it will depend on all the values in all the cells so it is it is a different type of approach it is quite useful for doing some frequency domain filtering high pass filtering which increases the sharpness or low pass filtering which smoothes the image that kind of thing can be done in the frequency domain okay now let us concentrate on our spatial domain (refer slide time35:55)



approach or operations okay so here we are going to look at one first thing we are going to look at smoothing okay smoothing is basically used for removal of noise or reduce the effect of reduce the effect of random noise

okay noise always is present you have because of either reflections or you can have some electronic noise creeping in the image is going to have some variation in intensity which could be of random type now there are different approaches available here different mask can be used so let us look at first one which we call as neighborhood average averaging okay the idea is that you take a template and we want to find out the average intensity in this neighborhood and say okay that should be the new intensity for the central pixel

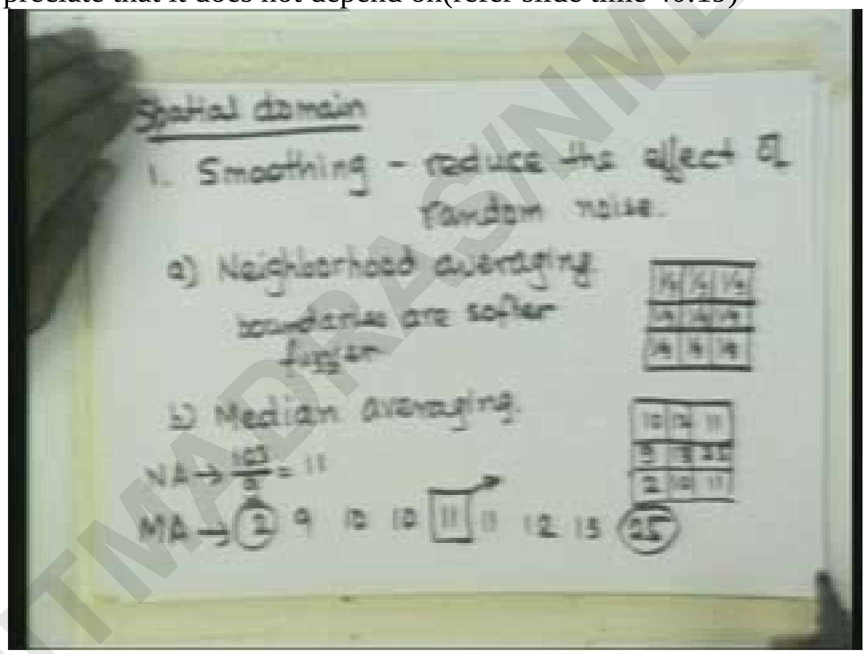
so if I take three by three template then the weights I should give this what I should give one by nine in each of these cells okay so automatically when I place this mask on different parts of the image then each of the intensities will get multiplied by one by nine and then we will sum it all together so it will give me the average intensity and then that will become the new intensity at the center pixel ok so this is what is the effect what is the other effect so what we doing is supposing there is some random noise is present we will hope to cancel it that by averaging it over a region region doesnot have to be three by three it can be it can be five by five seven by seven eleven by eleven whatever depends on how smooth you want the processed image to become now the disadvantage of that is that you are loosing some information what is going to happen is the sharp edges are going to get dull specially my mask is very large then I will have a very fuzzy kind of boundaries because as I approach the boundary the effect of the boundaries felt and that starts showing up in the processed image so this the disadvantage is that you get basically boundaries are softer or fuzzier okay

so you will lose information supposing you had some text written on it or some small mark which was significant then it will get lost so that is normally you would not do averaging for more more than five by five mask it also takes more processing if you have bigger mask the better smoothing sometimes is what is called median average okay in this case particular we cannot show this as a mask the idea basically is that if you have a part of the image supposing I am applying this to again three by three and I have some intensities in each of these cells say I have some image like some ten here twelve here eleven here and making one up thirteen here twenty five here two ten and say eleven this is one part of the image where I am going to apply both the masks if I applied the first mask

so in this case neighborhood averaging yields sum of all these which can be twenty two thirty three fourty two fourty five fifty five sixty eighty eighty two ninty two hundred three divided by nine so we have eleven roughly about eleven is the average that I am getting here okay but notice that it did get influence by this particular one so if this was not twenty five but hundred percent reason because there was some noise then this value will get also changed quite a bit will increase by a factor of by about ten right

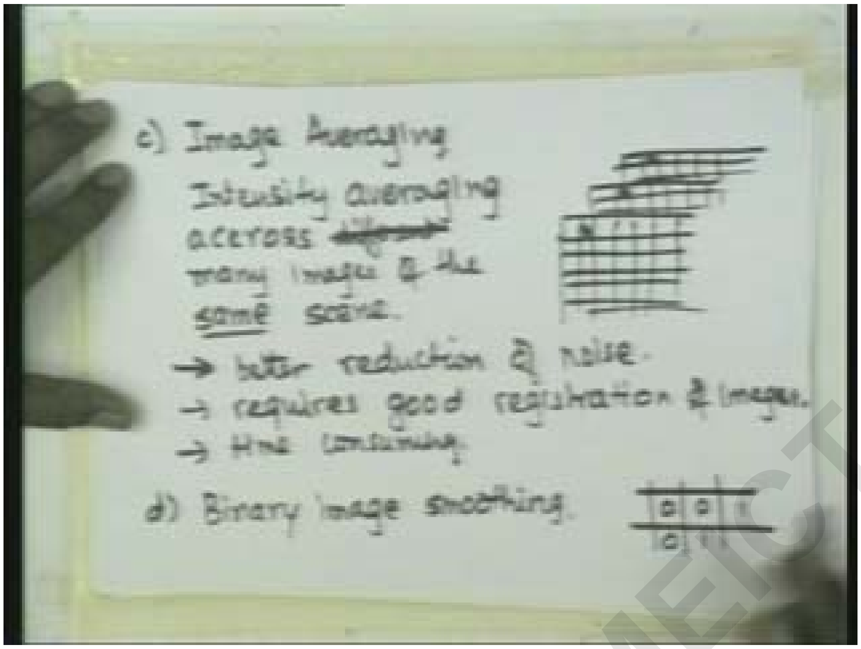
now if I use median averaging then what I do is I arrange these intensities in the increasing order so I have two followed by next intensities next intensities nine next intensity ten I have another ten then I have eleven I have another eleven then I have twelve thirteen then I have twenty five so median averaging what I do is I look at the then middle value after I have arranged it in a sequence

so it is the fifth one out of nine and so this is again turns out to be eleven here right but you will appreciate that it does not depend on(refer slide time 40:19)



what this value is or what this value is right so it makes it gives you immunity from outliners some value was very different then your average does not depend on that so it is in many case turns out to be superior

okay it requires a little different type of processing and cannot be simply be defined as a mask ok one approach is what is we looking at is noise reduction now smoothing process something called image averaging (refer slide time43:44)



okay here what we talking about is that we have a one image here with all the intensities then I have the same image then I have another image of the same scene I have another image of the same scene and just putting behind the other

so if I take all the corresponding pixels and find the average of the intensities then this is called image averaging so this is intensity averaging across different suppose the same image but many images is better than different images many images of the same scene right this way it is obvious that whatever average random noise is there if I take many of these images then I will be able to reduce the noise I will be able to get lesser variation in basically variation which is there because of random randomness is going to be reduced ofcourse it is a time consuming process you have to have several images to be able to extract out the information and also um yeah basically it is not that useful because if you want to do anything in real time then you dont have information available right away you have to accumulate several images and the fact that the images have to be registered you should not have had any motion from of any object in the image okay

so the advantage as I said is that you get better reduction of noise okay but requires good registration of images okay also time consuming okay final thing I would like to talk about in averaging or smoothing is dealing with binary images binary image smoothing okay in binary images the main difference is that each of the pixels is either zero or one so you can have whatever intensities are only limited to two values it is zero and one so the processing also becomes correspondingly simpler because we can process binary images using some kind of boolean expressions

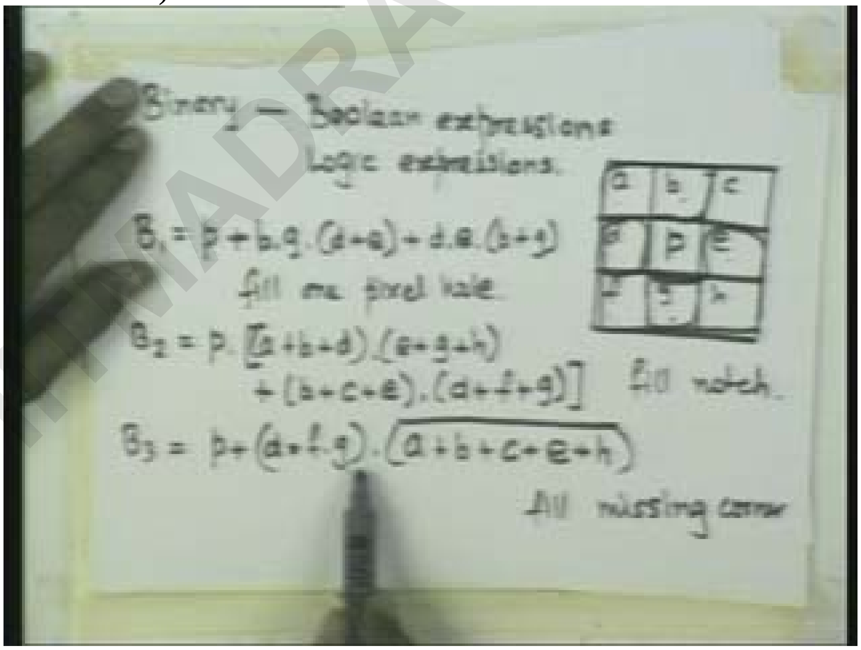
okay boolean expressions are essentially logic expressions okay so let me now instead of writing x minus one y minus one etcetera etcetera what I am going to do is I am going to do in this particular template let me label them as a b c d e f g h and the center one I am calling as p which is the pixel of which is being processed okay

so supposing I have some expression b one which says p or b and g and d or e or d and e and b or g right you would be familiar that I am calling or operation of the boolean as plus sign and dot as or multiplication as and operation so what does this mean this means that if I am looking at pixel p so either pixel p is one or b which is this one g which is this one and one of these two right if this is the one this is the one this is the one then even if p was not the one to begin with I should put one right because that is what this expression is saying either p itself is one if it is one then there is nothing else really matters because or it can only It cannot make it zero right

so it is true then p is true that means it is one then that is okay if it is not one if it is zero then it can be become a one if either these three were ones or these three were ones or alternately from the second term here these were ones or these three were ones so basically we are saying that If there is a missing pixel okay which is part part of some object because there are atleast three pixels here which are all ones so there is its not all noise

so therefore this could be made one so this is basically going to replace for us till one pixel hole okay similarly I can define another expression p and a or b or d and e or g or h or b or c or e and d or f or g okay

so here what I am saying is that if p was zero to begin with then It will remain zero if p was one to begin with then depending on a b or d so I am looking at this now this region if one of this atleast was five or true and one of this it is in the opposite side was high then this should be made a one otherwise it becomes like you know like may be there was a oblique line passing through here and so I have these pixels some of these pixels have one some of these pixels have one but this is zero which is possibly because of noise right so I am going to fill this off so it will fill a small notch one pixel one pixel one pixel notch likewise I can have another expression say p or d and f and g and a or b or c or e or h not of the whole okay now what is this to see if pixel p was one then it is okay it means one if it was not one it can become one provided one out of d f or g so basically one out of(refer slide time49:40)



sorry all of these not one out of but all of these were one okay and these were not one none of these were one and all three of these were one

so what is the operation you can characterize this by saying that its going to fill missing corners and this we have done for the upper right corner similar expression I can write for lower left corner or lower right corner or upper left corner okay by considering different set of pixels here so you get the idea that if you have a binary expression of course the idea is that the processing is very fast here we are only doing binary operations which are very much natural to the computers and also we dealing with pixels which are so processing can be very fast okay ofcourse amount of information is less here but processing can be very fast we can do some amount of filling of notches missing pixels corners and in general enhance the or make your processed image better than the original now what is better what is not better often timeful depend on the situation these are not universal formulas which will be useful everywhere but depends on what is the kind of problem the image is having you will have to design your filters which will take care of the particular problem and it may create some other problem

so it is a question of lot of experience is there which you have to utilize to come out with the right filter for the image so we will stop here