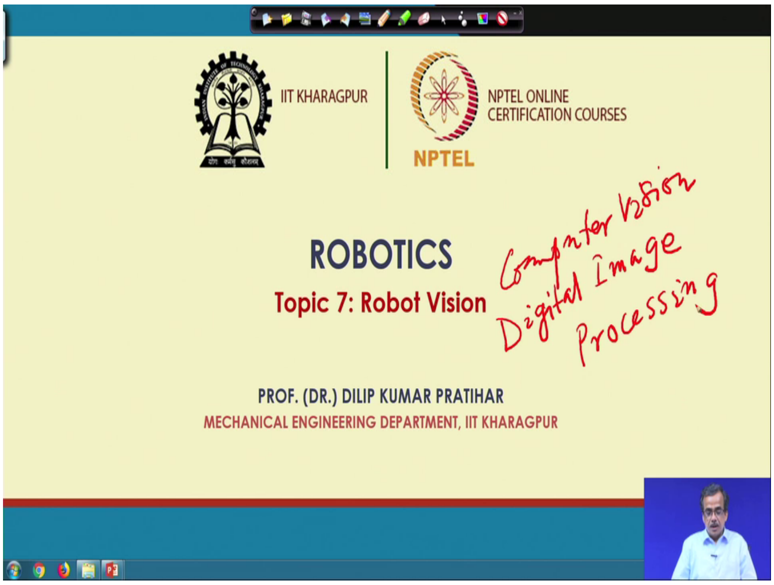
Robotics
Prof. Dilip Kumar Pratihar
Department of Mechanical Engineering
Indian Institute of Technology, Kharagpur

Lecture – 34
Robot Vision

We are going to start with a new topic that is a topic 7, it is on Robot Vision. Now, this robot vision this is also known as the computer vision computer vision.

(Refer Slide Time: 00:28)

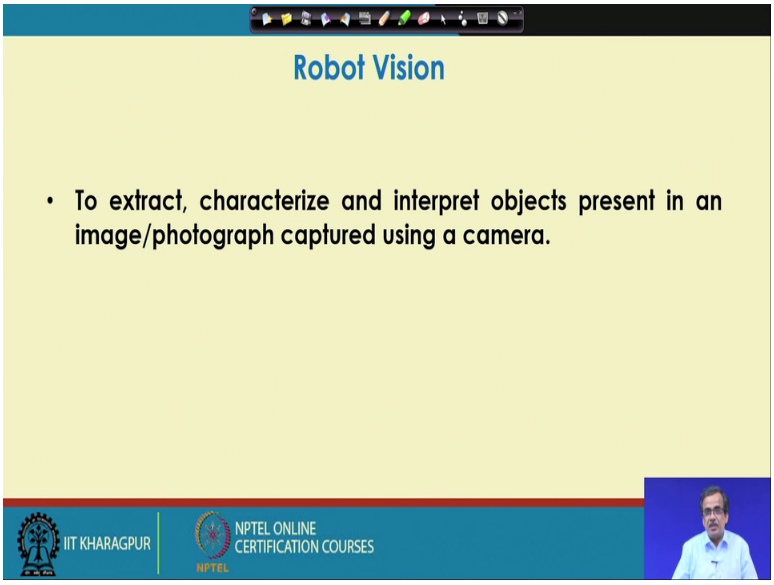


And it robot vision we use the principle of the digital image processing. So, we are going to use the principle of the digital image processing. Now, the aim of this robot vision, or the computer vision is to help the robot to collect information of this particular the environment.

Now, let us see how can a robot can collect information of the environment with the help of camera. Now, before I discuss further now let us try to see the way we human being do collect information of the environment, with the help of our eyes. So, with the help of our eyes we take photograph, or snap of the environment and there is a lot of processing in our brain and consequently, we could identify that this is object a, this is object b, present in a particular that the scenario, or project in a particular the image, or photograph.

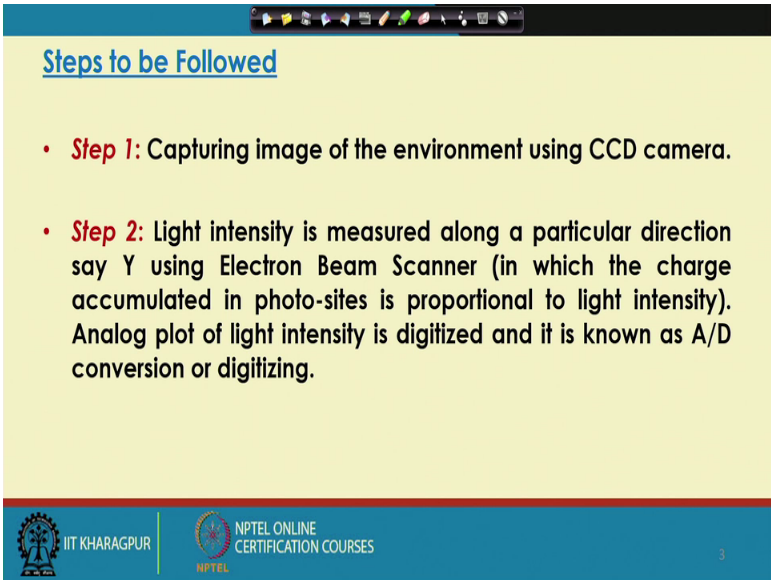
Now, exactly this particular principle we are going to copy in the artificial way in robot vision, or the computer vision. Now, once again let me repeat the purpose of robot vision is to identify and, interpret the different objects present in a particular image or the photograph. Now, let us see how to how to carry out. So, this particular the digital image processing, or the robot vision or the computer vision.

(Refer Slide Time: 02:23)



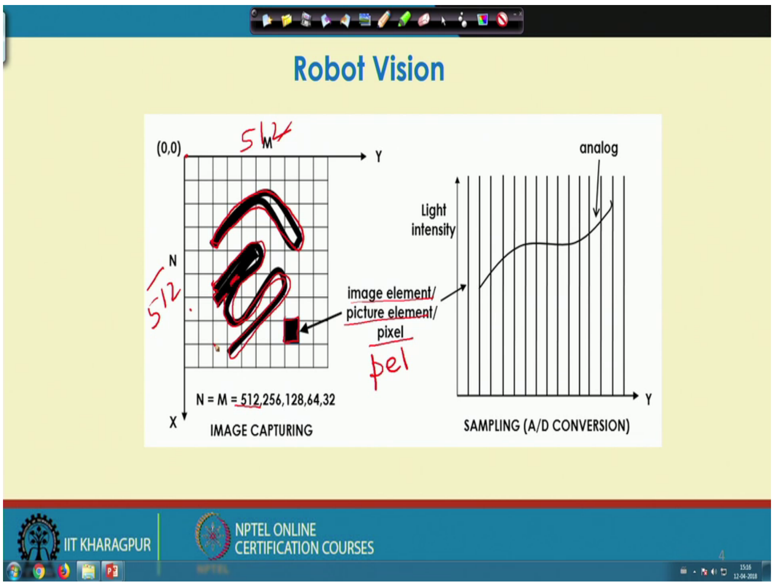
Now, the purpose as I told to extract characterize and interpret objects present in a scenario, or a photograph with the help of camera. Generally we use some sort of CCD camera the charged coupled device camera.

(Refer Slide Time: 02:43)



So, we capture image with the help of the CCD camera now as I told so, by CCD we means charge coupled device camera. So, this is actually the step one of this particular the computer vision or the robot vision and, once we have got this particular the image of the environment collected with the help of so, this particular camera so, it looks like this.

(Refer Slide Time: 03:10)



Supposing that so, this is the computer screen. So, this is the positive Y direction and this is the positive X direction. Now, here actually what I do is so, with the help of camera the image or the environment, whose photograph we have taken. So, that is actually transferred to the computer and on the computer screen will be able to see this type of image. Now, what I do is so, this particular the computer screen that is divided into a large number of small small segments for example, say along Y direction. So, we take m number of division along X direction, we take N number of divisions and this is nothing, but the origin that is 0 0.

Now, if this is the computer screen so, that is divided into M cross N. So, so many such small small subdivisions for example, say so, I will be getting M equals to N equals to 512 or 256 or 128 or 64 or 32. Now if I take M equals to N equals to 512 that means, here there are 512 divisions and here, also along this particular X direction there will be 52 are divisions.

That means so; this particular area is divided into like 512 multiplied by 512. So, so many such small small image elements and this image element for example, say this is one small image element and this is known as the image element, or the picture element or the pixel or in some of the literature this is also known as pel.

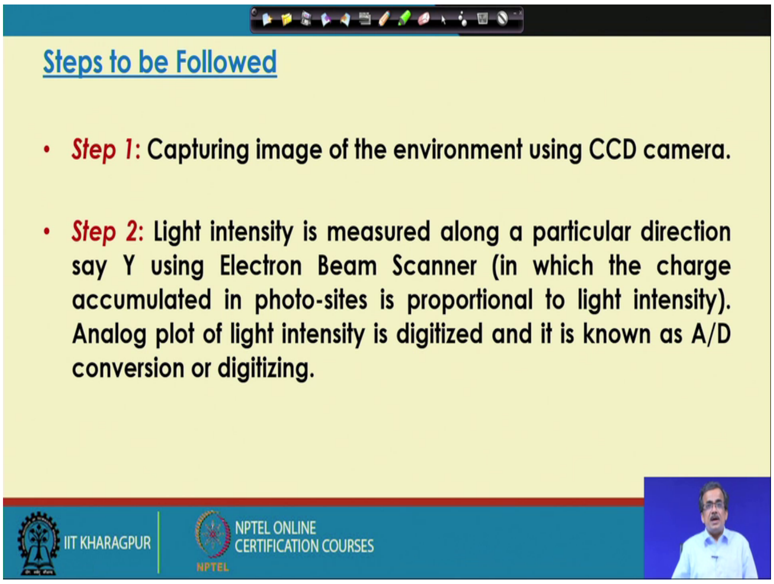
So, this particular pixel so, so many such pixel we have, now will have to concentrate on this particular the pixel. Now, supposing that with the help of camera and so, we have got this type of image so, for example, say this is one image which I have gotten this particular computer screen and this is collected with the help of say camera and that is this your transferred to the display of this particular your the computer ok..

Now so, if this is black and white picture, now the difference between the black and white is actually the amount of the light intensity for example, if I consider that this is the black object means, the light intensity is less and on the white portion the light intensity will be more.

Now, depending on this particular light intensity and the difference in light intensity, we can I identify the black and white for example, say if we take the photograph the black and white photograph of a human being. The head portion the hair portion will be black and the face will be slightly white is so, if I compare the light intensity values of the hair and the face, the light intensity value of the face will be more compared to that of the your the hair part that is the black hair part. So, this is the way actually we can find out the difference between the black object and white object due to the difference in light intensity.

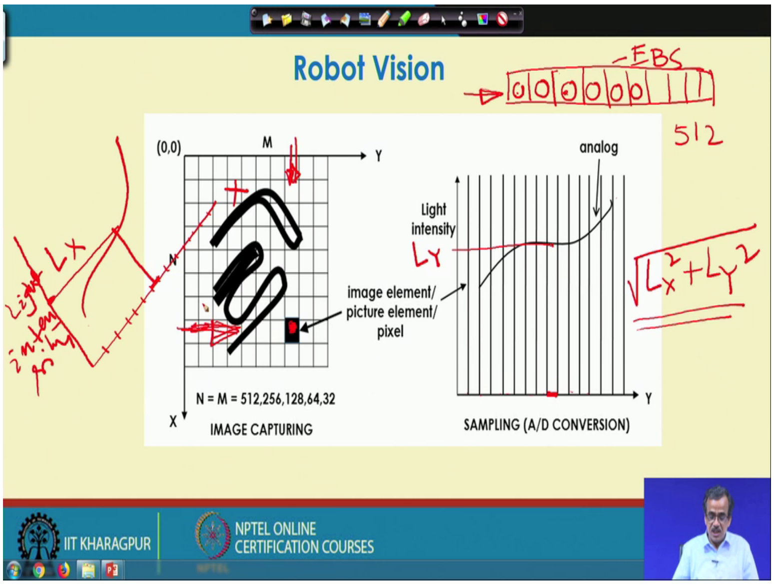
Now, supposing that we have got this particular picture and now actually we are going to find out, what should be the light intensity at each of these particular the pixel. Now to do that actually what I do is we try to take the help of step 2.

(Refer Slide Time: 07:19)



Now, this particular step two is nothing, but actually your we take the help of one electron beam scanner and, we do the scanning along the Y direction and this particular the X direction just to collect the light intensity values at each of this particular the pixel.

(Refer Slide Time: 07:36)



Now, let us say how to do it now what I do is say we are just doing scanning in the positive Y direction supposing that I want to find out what should be the light intensity. So, at this particular the pixel so, what I do is we try to actually do the scanning along this particular the Y direction and, how to do this particular scanning, we take the help of one electron beam scanner. Now, this electron beam scanner is something like this, now here on this electron beam scanner there are some photo sides.

So, we have got a large number of photo sides here for example, if there are 512 division so, I can consider the 512 photo sides. And these particular electron beam scanners this is electron beam scanner and that is put so, just say below that and we do the scanning the moment with do this is this particular scanning along this particular Y direction. So, what will happen is so, due to the variation of this light intensity. So, different amount of electrical charges will be accumulated on this particular the photo sides for example, say here I have got a photo side, I have got another photo side here another photo side. So, these are all photo sides and now I am doing the scanning in this particular the direction.

Now, if the light intensity is more, the more amount of charge will be accumulated in the photo side, on the other hand if the light intensity is less; that means, I am passing through the black region so, less amount of charge will be accumulated in this particular the photo side, now what I can do is we can measure how much is the amount of accumulated this particular the charge in each of the photo sides. And here we just prepare one plot and this particular plot is nothing, but light intensity verses the Y direction, now here each of these particular indicates actually the your the pixels.

So, starting from here so, pixel wise I can plot; that means, along this particular Y ok. So, I can plot what is the variation of this particular the light intensity value and, this particular information is nothing, but the analogue information. Now, this is what is happening along the Y direction ok. And supposing that I am concentrating here so, corresponding to this particular pixel. So, I am getting that this is the amount of your the light intensity and supposing that that is denoted by say L y. The same thing we do along this particular the X direction. So, what I do is we try do the scanning in this particular direction, in the positive X direction and once again we will pass through the same pixel ok.

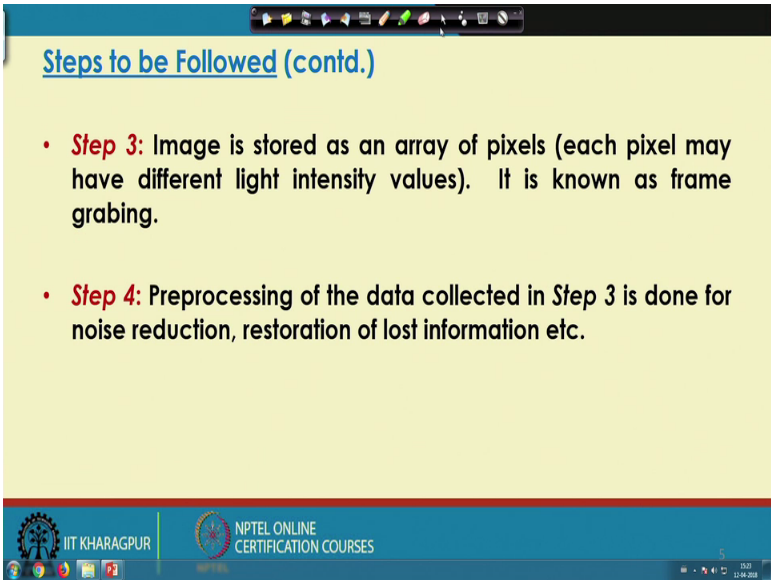
And exactly in the same way if I just plot, supposing that this is the light intensity say for example, say this is the light intensity light intensity and this is the X direction. So, once again I have got all such your the pixels. So, for each of these particular pixels. So, I will be able to find out what should be the analogue plot for this light intensity..

For example, say I am I am just moving along this particular the positive X direction. So, there is every possibility that I will be getting some sort of your the profile of light intensity like this. And once again if I concentrate on the same pixel supposing that I am here. So, I will try to find out what is the light intensity value corresponding to that particular the pixel and supposing that so, that particular numerical value is nothing, but L X.

Now, corresponding to this particular pixel so, I have got this L Y and this particular L X. So, after that actually what I do is we try to find out what is this square root of L X square plus your L Y square. So, I will be getting some the numerical value. So, I will be getting some real value and we try to find out what is the nearest integer. Now, corresponding to that corresponding to this value, corresponding to this the nearest integer will be the light intensity value, corresponding to this particular the pixel. The same process we follow for each of these particular the pixel.

So, we can find out what should be the light intensity value ok, corresponding to each of these particular the pixel. Now, here so corresponding to this particular image, so, I have got some sort of the light intensity values ok.

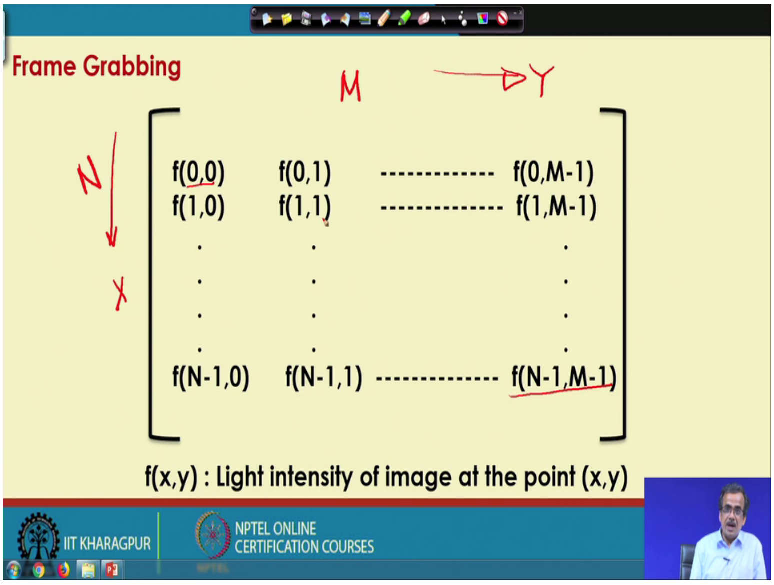
(Refer Slide Time: 13:04)



Now, the step three whatever I mentioned so, that is nothing, but the step 3 that is image is stored as an era of pixel. And at each pixel actually we try to mention what is the light intensity value and this particular process of storing an image, or a photograph with the help of some numerical values of light intensity. So, this is what is known as the frame grabbing.

In fact, your unless we do the frame grabbing so, will not be able to carry out any such your the calculation with the help of this computer. Now, because computer does not know anything except the numbers so, what I will have to do is corresponding to that particular image. So, I will have to find out the corresponding your the matrix the matrix of light intensity values. And this particular process is known as your the frame grabbing. And once that particular the frame grabbing is done.

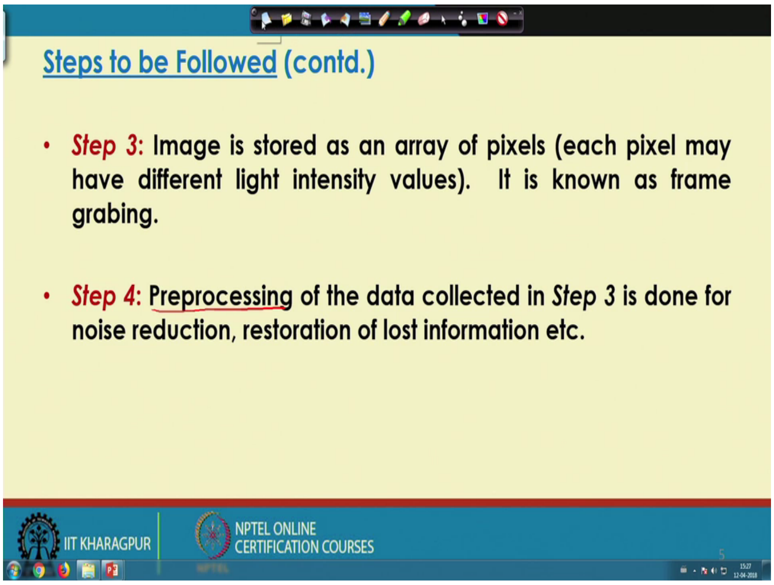
(Refer Slide Time: 14:16)



Now, actually we are in a position to represent so, this particular image in the form so, this type of matrix. Now, here if you see for example, say we consider there are m number of M number of divisions along this particular positive Y direction. So, this is the positive Y direction and, we consider capital M number of division. Now, this is actually the positive X direction and we consider capital N number of divisions ok. Now, here so, this f 0 0 is nothing, but the light intensity corresponding to the pixel whose coordinate is 0 comma 0. Similarly so, this f N minus 1 M minus 1 is nothing, but the light intensity value corresponding to the pixel whose coordinate is N minus 1, comma M minus 1.

So, for each of this particular pixel so, we can find out the light intensity values and this values are nothing, but is your integer values. Now, here I have written so, f x y indicates the light intensity of the image at the point x y. So, similarly this f 1 1 is a light intensity value at the point whose coordinate is your 1 comma 1. So, this is the way actually we can represent one image with the help of a matrix of some numerical values and, this numerical values are nothing, but the light intensity values in the integer form. Now, let us see like how to how to proceed further.

(Refer Slide Time: 16:06)



Now, here actually what will have to do is the if you see this particular the matrix of light intensity values. So, this particular the matrix cannot be very accurate, or the correct the reason is very simple. The quality of this particular image, or the light intensity values depends on a number of parameters for example, it depends on the level of illumination at which we are taking that particular we are collecting that particular picture.

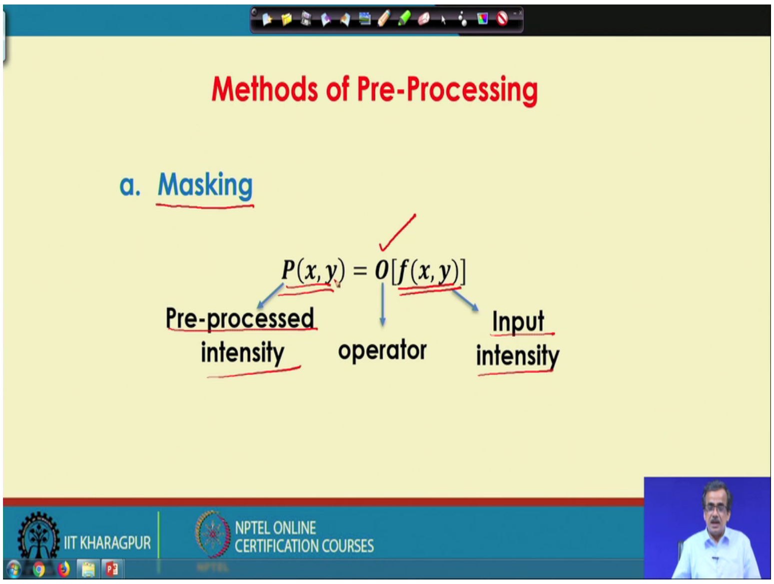
It depends on the angle at which I am collecting that particular picture; it depends on actually the expertise of the operator. So, this data which you have got corresponding to this particular image may not be very accurate. So, there could be some noise, there could be some sort of in imprecision, there could be some sort of uncertainty and that is why we take the help of one step that is called the preprocessing.

So, we try to do some sort of the pre processing and, if you do this preprocessing actually we can remove this particular the noise from this the matrix. So, the purpose of preprocessing is to the remove noise, from this particular the light intensity values, or sometimes there is a possibility that some part of information will be lost from this particular picture and, we try to restore that particular the information.

So, if you want to reduce the noise from this particular data, or if you want to restore some sort of lost information, will have to take that help of some sort of preprocessing. Now, if you see the literature in fact, we have got different methods for this particular the processing.

Now, here I am just going to discuss the principle of each of these particular the processing methods. So, this is actually the thing which will be getting so, this particular matrix corresponds to your the image.

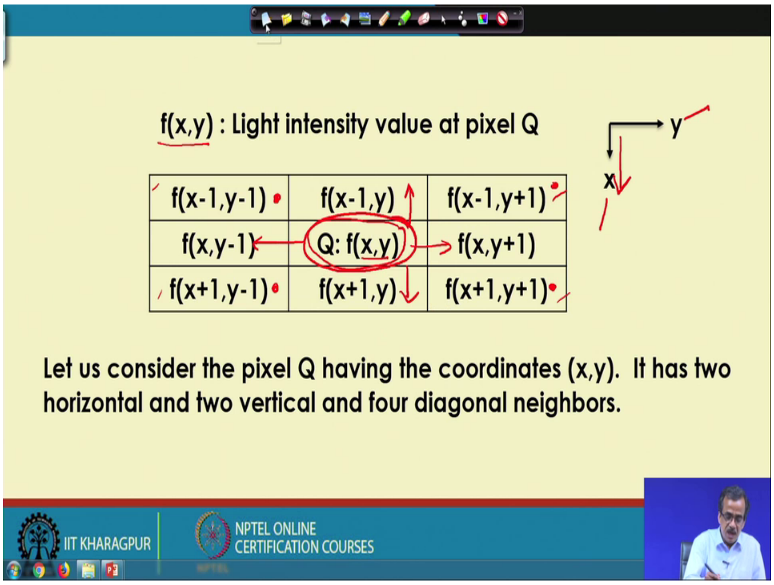
(Refer Slide Time: 18:24)



Now, let us see how to do this particular the preprocessing just to reduce that noise from the data. So, methods of preprocessing as I told there are several methods and out of all such methods, I am just going to discuss a few very popular method for example, say the masking is a very popular method for preprocessing. Now, the method of masking is a very simple actually what I do is supposing that so, this f x y. So, this is nothing, but the light intensity value at the pixel, whose coordinate is x comma y.

And on this particular light intensity value, we use one operator that is nothing, but O. So, this operator O is going to work on so, this f x y that is the light intensity value which is nothing, but the input intensity. And we are going to find out what is this preprocessed intensity that is P x y. So, our aim is to determine so, this particular the P x y. Now, let us see like how to how to determine this particular the P x y.

(Refer Slide Time: 19:44)

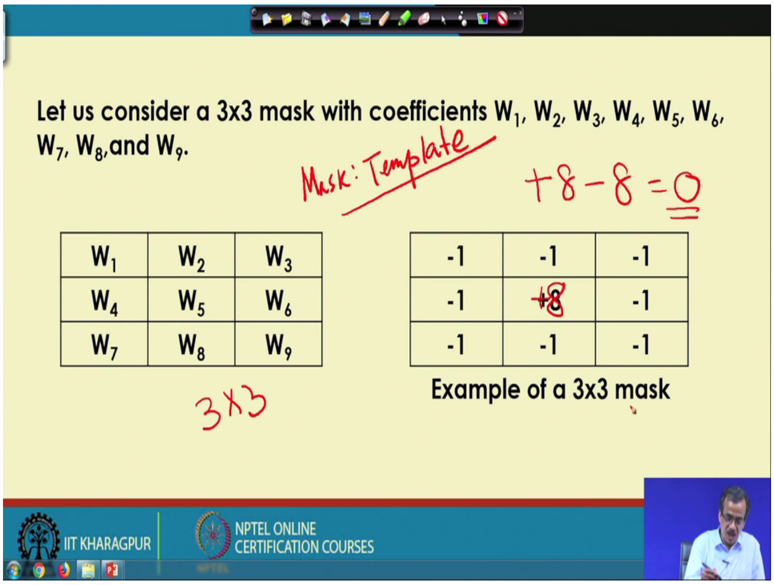


That is the preprocessed data. Now here, I am just going to concentrate on a particular pixel and it is neighbourhood. Now, here so as I told that f X Y is going to indicate the light intensity value at the pixel whose coordinate is nothing, but is your x y.

Now, this is the positive direction of Y this is the positive direction of X. So, starting from here. So, if I move along this particular direction. So, Y is going to increase so, this will be your X comma Y plus 1. Similarly here this will be X comma Y minus 1, because this is in the negative direction of Y. Now, similarly starting from here if I just go down, then what will happen is your so, this particular X is going to increase, because this is the positive direction of X. So, this will become f of X plus 1 comma Y and the coordinate of this particular pixel will be your X minus 1 Y.

Similarly, I can also find out the coordinate of this and this. Now, if I concentrate on this particular pixel, that is denoted by Q whose coordinate is X Y and whose light intensity is f X Y. Now, this particular the pixel has got two horizontal neighbor it has got two vertical neighbour and, it has got four such diagonal neighbours. So, once again let me repeat let me repeat that for a particular pixel, there are two horizontal neighbours two vertical neighbours and there and there are four such your diagonal neighbours. So, will have to concentrate on this particular horizontal vertical and the diagonal neighbours ok.

(Refer Slide Time: 21:58)



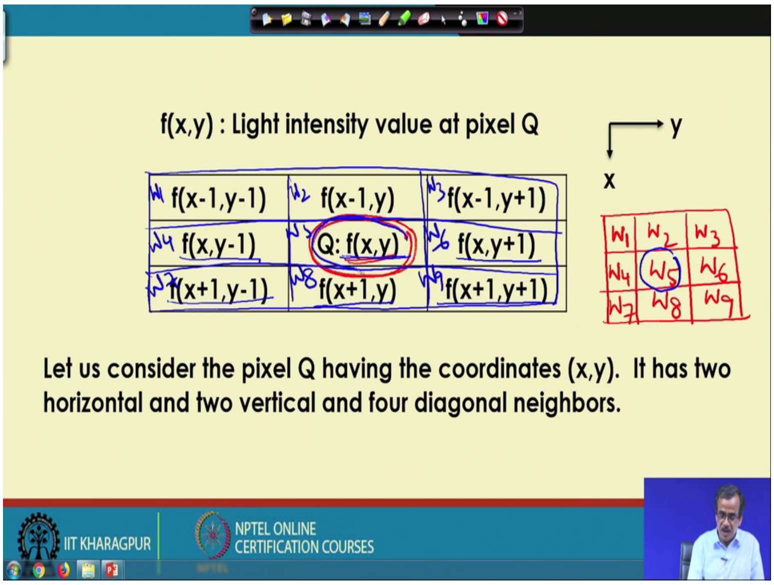
Now, let us see how to carry out so, this particular the pre processing. Now, here actually in masking what we do is we try to help take the help of a mask and this particular mask is nothing, but a template. So, by mask we mean so, this is nothing but a template, now with the help of this particular template actually we can do this masking operation, or we can do this particular the preprocessing..

Now, here this shows a typical 3 cross 3 mask and this W values are nothing, but the coefficient of the mask for example, W 1, W 2, W 3 up to say W 9. So, this is a 3 cross 3 mask so, there are 9 such W values and these are nothing, but the coefficient of this particular the mask ok.

Now, here it shows a typical 3 cross 3 mask, now here we can see that. So, here I have put plus 8 so, this is plus 8 and here you can see I have put minus 1 minus 1 minus 1 minus 1, here minus 1 minus 1 minus 1 minus 1. So, if we just add all such minus 1 values. So, I will be getting 3 plus 2 5 plus plus 3 that is 8. So, I have got plus 8 minus 8 that is your is equal to 0.

So, the sum of all this particular coefficient values will be equal to 0. So, this mask coefficient values are selected in such a way. So, that the sum of this particular mask coefficient values becomes equal to 0. So, as I told that this is one typical 3 cross 3 mask which is very frequently used for the preprocessing. Now, let us see how to implement.

(Refer Slide Time: 24:03)



So, this particular the masking operation, our aim is to find out supposing that. So, this is actually your one say image and these are actually the light intensity values at the different pixel and, our aim is to find out what should be the preprocessed value, corresponding to this particular f X Y. And if I want to find out what should be the corresponding preprocessed value, for this particular f X Y, what we do is we try to take the help of one template, or the mask and let be considered 1 3 cross 3 mask, or 3 cross 3 templates. And as I discuss so, the coefficients are W 1; W 2, W 3 then comes W 4 W 5, W 6, then comes W 7 W 8, W 9. So, these are nothing, but the mask coefficients and how to find out to the pre processed value corresponding to this.

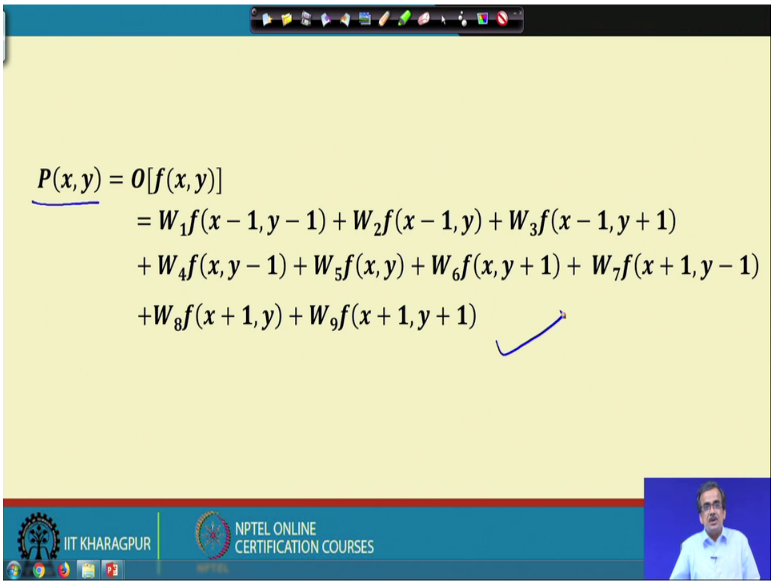
The method is very simple actually, what we do is actually, what I do is so, this is this is actually its corresponding preprocessed value I will have to find out and, you concentrate on the mask centre supposing that this is the mask center. So, this particular template or the mask you bring it here and this particular mask centre is going to coincide with this particular your this pixel.

So, what I am going to do is I am just going to put this particular mask here. So, as if I am just going to put the mask something like this and here, I am just going to write down all such you are the mask coefficient W 1, W 2, W 3, then comes your W 4, W 5, W 6 then W 7, W 8 and W 9 and after that actually what we do is.

So, we multiply so, this particular W 1 with f of X minus 1 Y minus 1 plus W 2 multiplied by f of X minus 1 comma Y plus W 3 multiplied by so, this particular light intensity value plus W 4 multiplied by this light intensity value, plus W 5 multiplied by this plus W 6 multiplied by this, W 7 multiplied by this f W 8 multiplied by this f plus W 9 multiplied by this f we sum them up..

And then will be getting some numerical value and that particular numerical value is nothing, but the preprocessed value corresponding to this particular the light intensity value. So, this is the way actually we can find out like, what should be the preprocessed value corresponding to that particular your the pixel.

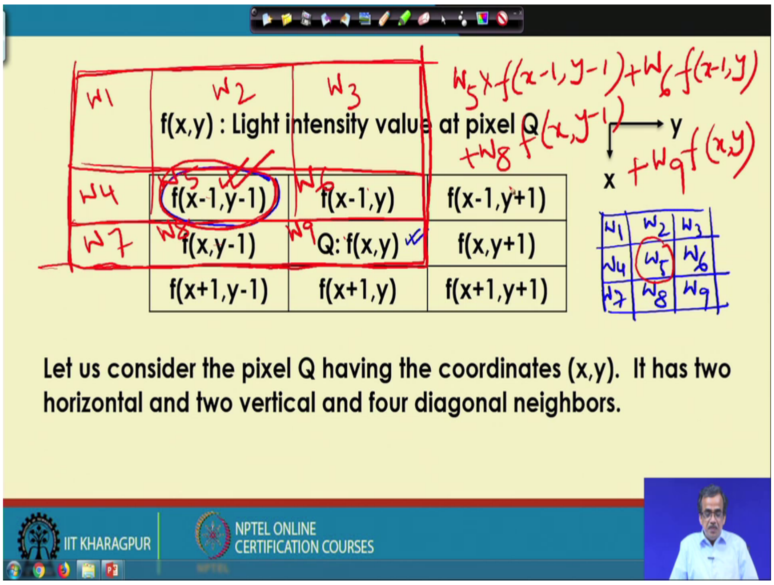
(Refer Slide Time: 27:21)



Now, the same thing whatever I discussed the same thing I have just written it here. So, P x y is nothing, but the operator o that is acting on f x y and if we remember. So, W 1 multiplied by this f W 2 multiplied by this f so, whatever I discuss the same thing I have written it here. So, this is the way actually we can find out the preprocessed value, corresponding to that particular the masking.

Now, here I am just going to take another very small example, like how to determine the preprocessed value corresponding to this particular your corresponding to this particular the light intensity value. So, the we have already seen how to determine the preprocessed value here, but let us discuss how to determine the preprocessed value corresponding to this particular the pixel whose light intensity value is nothing, but f of X minus 1 comma Y minus 1.

(Refer Slide Time: 28:32)



Now, as I told that will have to take the help of your the mask that is nothing, but the 3 cross 3 matrix. So, this is W 1 W 2 W 3 W 4 W 5 W 6 W 7 W 8 and W 9. So, our aim is to determine actually what should be the preprocessed value corresponding to this particular the pixel. So, what we do is we concentrate on this particular the mask center.

So, once again this is the mask center so, this particular mask center is made coincident with this particular the pixel; that means, W 5 will come here and so, this will be the W 8 sort of thing. So, I will have to put the mask something like this. So, this is the way I can put this particular mask, so here so, this particular W 5 will be here so, I will have to do something like this ok.

And now we can see that we have got so, this is your so, this is the way actually we can do, so we put this particular mask here. So, this is the mask which I am going to put. So, this is this corresponds to your W 1, this is your W 2, this is W 3, this is your W 4 and here you have got W 5 and we have got W 6 here, then comes your W 7 here, then comes W 8 here and this is your W 9 and our aim is to find out the pre processed value corresponding to this.

Now, if this is the scenario so, the contribution of this particular W 1 W 2 W 3 W 4 and W 7 will be equal to 0 here ok. So, now, will have to concentrate only on so, this particular 1 2 3 4. So, only this four will have to concentrate. Now, if we concentrate only on this particular the 4. So, I will be able to find out the preprocessed value is nothing, but W 5 multiplied by f of x minus 1 y minus 1.

That is this particular thing plus W 6 multiplied by f of x minus 1 y plus W 8 multiplied by f of x y minus 1 plus W 9 multiplied by f of x comma y. So, we can find out the pre processed value corresponding to this particular the pixel, the same procedure actually I can I can follow at each of the pixel, just to find out the preprocessed value corresponding to that particular the pixel.

Thank you.