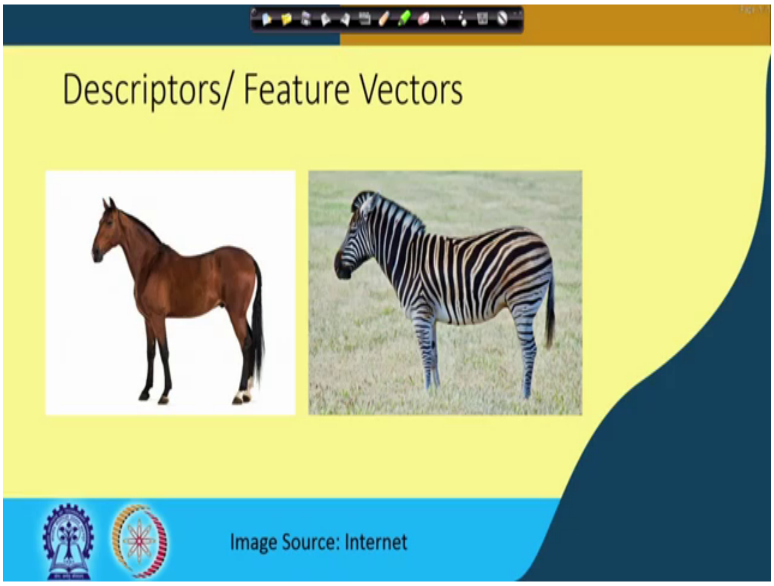
# Deep Learning

**Prof. Prabir Kumar Biswas**
**Department of Electronics And Electrical Communication Engineering**
**Indian Institute of Technology, Kharagpur**

## Lecture - 02
## Feature Descriptor - I

Hello, welcome to the NPTEL online certification course on Deep Learning. In today's lecture, we will try to find out how do we capture the information or the descriptions or from the signals that we capture from the real world.

(Refer Slide Time: 00:51)




In our previous introductory lecture we had given, we have shown you the images of a horse and a zebra and then we have told that to differentiate between these two images that is to identify that which is the horse and which is the zebra, we can extract two types of descriptors one is the shape descriptor that is the shape of the horse and the shape of the zebra and the second one is the region descriptor which means that, what is the content of or what is the color intensity and texture of the body of the horse and the body of the zebra.

So, if you look at the shape of these two animals the horse and the zebra, you find that the shape boundary is more or less same; that means, this shape information or the boundary information does not give you the sufficient description or sufficient information by which you can differentiate between a horse and a zebra.

But when you look at the entire figure that is along with the shape information when we also consider the information of the color, the information of intensity, the information of texture then only I can differentiate or I can see that which is the horse and which is the zebra.

So, that makes two things very clear that for images or for objects that we see in the real world, we can have two types of information one is the shape information or the boundary information and the other kind of information is the region information which gives you what is the intensity or what is the color or what is the texture. And when we combine both these informations that is the shape information, the color information, the intensity information and the texture information all of them together can identify a particular object or a particular animal.

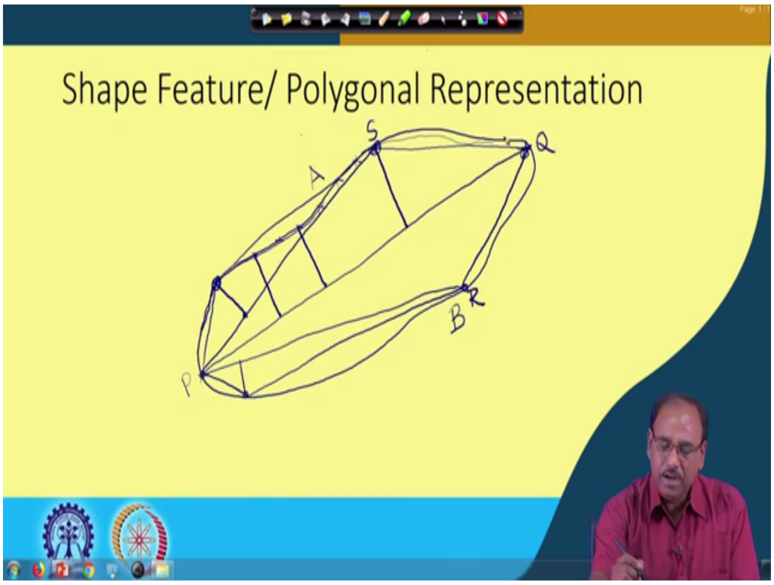
(Refer Slide Time: 03:05)


So, in today's lecture, what we are going to talk about that how we can obtain the descriptors or features from the signals that we obtain from the real world. These signals can be visual signals as in the form of images or what we can see through our eyes. These signals can also be audio signals that we can hear and signals like the speech signals or voice signals using which I can differentiate among different speakers I can also understand what is being spoken.

The applications can be speaker identification, speech to text conversion and many such applications. So, firstly I will talk about the visual signals or how do I extract the different types of features or different types of x descriptors from a visual signal and these descriptors will can be of two types as we have already said that the descriptors can be obtained from the boundary which tells you what is the boundary property. It can also be obtained from the region which tells you what is the region property that includes the intensity that includes the color and that includes the texture properties.

So, firstly, let us see that how we can obtain the different types of boundary features or how we can extract the boundary properties.

(Refer Slide Time: 04:43)




So, for that I will take a very simple shape. So, I take a boundary shape something of this form let me change the size of the.

Student: (Refer Time: 05:17).

So, suppose I have a shape of this form, which is a closed boundary. Now, find that though this closed boundary I have shown it as a continuous curve, but you remember that we are talking about the discrete signals or digital signals. So, this curve is not really a closed curve or a continuous curve rather this curve is consists of a set of discrete points. So, something like this I have a set of points on this district curve.

So, one of the ways in which such an arbitrary boundary can be represented is in the form of a polygon. The way we can represent an arbitrary shape at boundary in the form of a polygon is that we can recursively subdivide that arbitrary shape into a number of segments. So, the way we can do this is suppose I take two points on this boundary which are at maximum distance.

So, the first division will be that I draw a straight line passing to these points on the boundary which are maxima at maximum distance. So, once I do that this particular chord or this particular straight line subdivides this boundary into two sub boundaries one sub boundaries in the upper part so let me call it sub boundary A and the other one is on the lower part, let me call it as sub boundary B.

So, at the next level, we can again subdivide both these sub boundaries and for that we have to use a criteria. The criteria can be that I compute the perpendicular distance of different points on these boundary segments from the straight line. So, I can find out given a point over here, I can try to compute what is the perpendicular distance of this point on this boundary. I can also compute what is the perpendicular distance of this point on the boundary from the straight line.

And I find out a point on this boundary which is at maximum distance from this line segment. So, I can identify that maybe this is a point on the boundary which is at maximum distance. So, once I identify this point then this point subdivides this upper sub boundary into two parts.

So, one part is starting from this point to this point, the other part is starting from this point to this point. So, let me name this points the initial points let me name this as P and Q, this point is S. So, the next subdivision after doing the next subdivision, the vertices of the polygon that I get is something like this. Similarly, coming to the lower part of the boundary maybe this is a point which is at maximum distance from the straight line PQ and in that case this point let me call it point R.

So, this point R subdivides this lower part of the boundary segment PQ into two more segments one is PR, the other one is RQ. So, the next level of polygonal representation is by this. So, you will find that up to this I have got a polygon P, S, Q, R. So, this is the polygonal representation at this level of this arbitrary shape at boundary that we had.

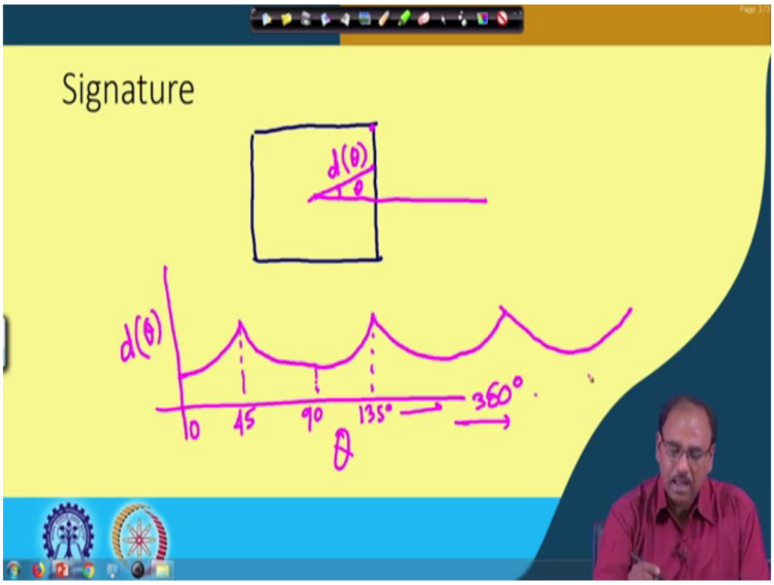
So, this process can be repeated recursively, this may be another vertex on this boundary because its distance from the straight line over here is maximum of the distances of all the points on this boundary segment from the straight line. So, this may be the next vertex, similarly some point over here may be in the next vertex and so on. So, each of these vertices if you join together, what I get is a polygonal representation of this arbitrary shaped boundary.

So, once I get such a polygonal representation, from this polygon I can try to extract different types of boundary features, simply from the properties of the polygon. So, this is one way of polygonization of an arbitrary boundary and once I have such polygonal approximation representation from this polygonal representation, I can obtain different types of descriptors or I can also obtain different types of descriptors from these boundary segments right.

So, here you find that I have different boundary segments starting from this vertex to this vertex this is one boundary segment, from this vertex to this vertex I have another boundary segment, from this vertex to this vertex I have another boundary segment. So, I can have different techniques to find out what is the shape of these different boundary segments.

So, the properties of the polygons as well as the properties of these boundary segments can give me important information about the shape of the boundary. So, this is one kind of descriptor or one kind of features that we can obtain.

(Refer Slide Time: 11:31)



The next kind of features that we can obtain is what is known as signature. So, to see what is the signature let me take a very simple shape which is a square.

So, I take a square something like this, what is signature? Signature is the plot of the distance of different boundary points from the centroid of this shape taken in various directions or in various orientations. So, as I have to compute or I have to find out the distance of different boundary points from the centroid of the shape in different orientations. So, I have to have a reference line. So, let me assume that my reference line is this.

So, this is the centroid and my reference line is this one. What I do is I compute the distance of the boundary point from the centroid which is oriented at an angle theta from the reference line and suppose this distance I call as d theta. So, what I plot is, the distance d theta against theta as theta varies from 0 to 360 degree.

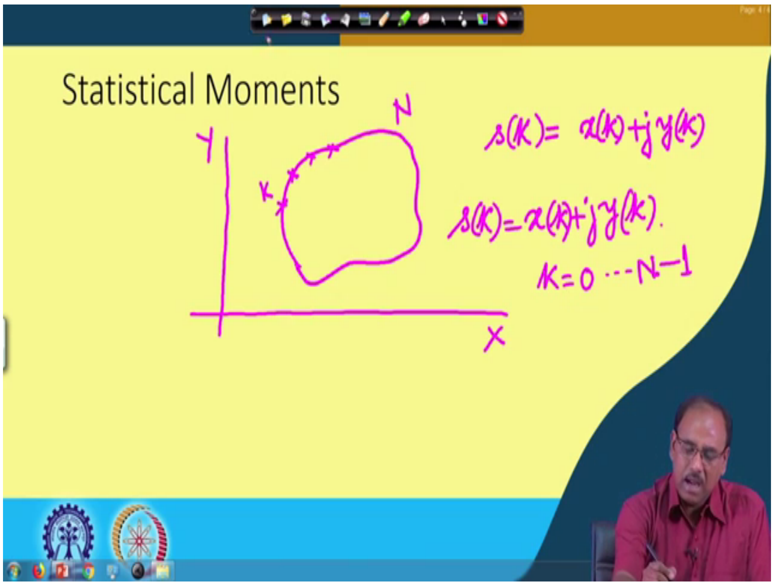
And in this case you will find that when theta is 0 as my reference line over is over here; the distance will be one of the minimals. So, I start from here and when theta is 45 degree as I have taken a square when theta is 45 degree; that means, I am computing the distance of this vertex of the square from the centroid this distance will be maximum.

So, I will have a d theta versus theta plot which will be something like this and it will go on up to 360 degree. So, I will have a minimized 0, I will have a maxima at 45 degrees, I will have another minima at 90 degree, I will have a maxima at 135 degrees and so on. So, such a plot of d theta versus theta gives you inform important information about the nature of the boundary or the shape of the boundary.

So, this is a kind of plot which I am very getting for a square figure if the figure instead of being square it is something else then naturally, naturally the nature of the plot will also be different. So, the shape of the plot or the nature of the plot also gives me important information about the shape of the boundary. So, by processing this signature I can also obtain boundary descriptors.

So, this is another way of obtaining the boundary descriptors the next type of features that we can obtain is what is known as Fourier descriptor. So, for to obtain the Fourier descriptor, I hope all of you know what is a Fourier transformation or Fourier coefficients. So, to obtain Fourier descriptor, I represent the boundary points the points on the boundary as a sequence of say complex numbers; let us see how we do it.

(Refer Slide Time: 15:38)

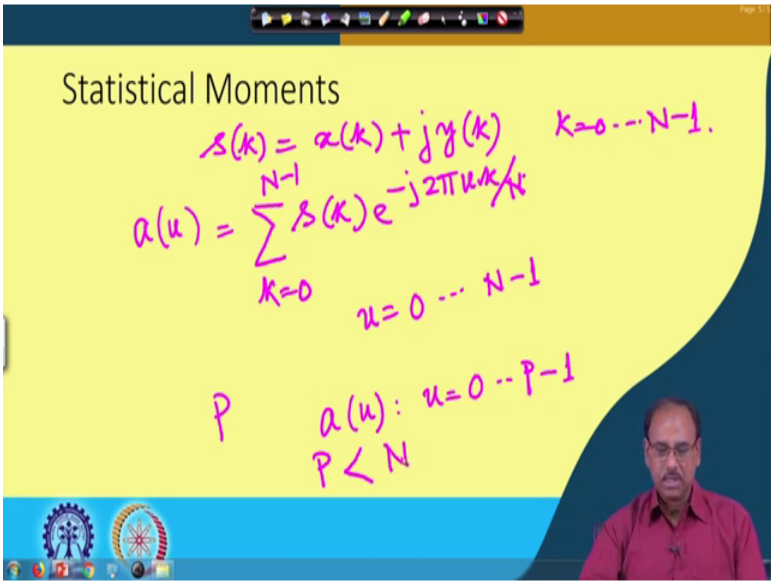



So, I assume two dimensional space and suppose I have a closed boundary something like this. As I said before that this boundary consists of a number of discrete points. So, I take any point say kth point on this boundary segment and this kth point I call it say sk which you have two components, one is in the X dimension direction and other one is in the Y direction.

So, this k will have two components which is xk and I represent this as say complex number. So, j times yk. So, when I consider all these different points on this boundary what I get is a sequence of such complex numbers.

So, you remember that all these points are to be represented as a sequence. So, I get a sequence of complex numbers which is given by this sk is equal to xk plus j times y k where k may vary from say 0 to capital N minus 1 where I have total number of points on this boundary which is capital N. So, once I have this, what I can do is this sequence of numbers can be represented by their Fourier coefficients.

(Refer Slide Time: 17:35)



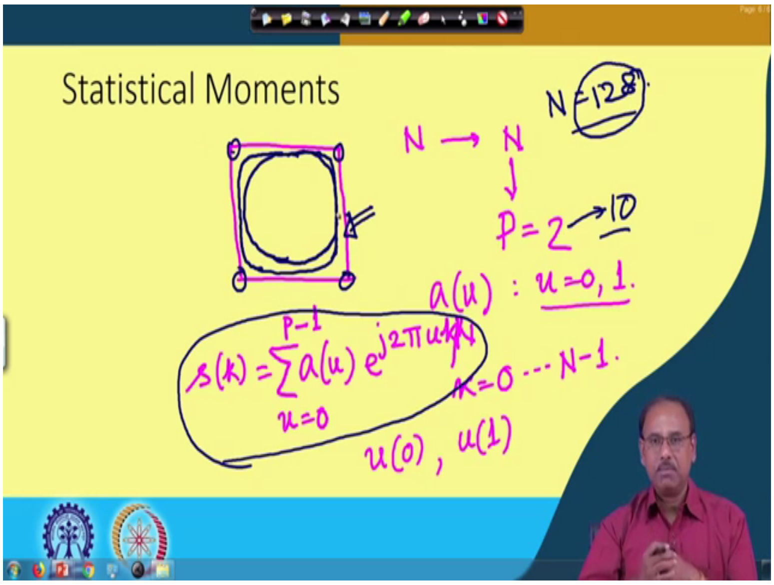

So, for that what I have to do is, I have to take the Fourier transformation of this sequence of numbers. So, I had sequence of numbers sk given by xk plus j times yk k varying from 0 to capital N minus 1 and what I do is, I take the Fourier transformation and you know the Fourier transformation is given by a u is equal to sk e to the power minus j 2 pi u k by n right.

I take the sum of this over k is equal to 0 to capital N minus 1 and this I will get coefficients I will get for all values of u varying from 0 to capital N minus 1. So, as I have capital N number of points in my sequence, I will also have capital N number of coefficients and the magnitudes of these coefficients represent or give me useful information of the shape of the boundary.

Now, you can also find out that instead of considering all the n number of coefficients if I consider say lesser number of coefficients then what kind of effect I can have. Suppose, I want to consider p number of coefficients; that means, I want a u where u varies from 0 to capital P minus 1 where P is less than capital N then what kind of effect this truncation of the coefficients you will have.

So, to see to see this truncation of the effect of the truncation of the coefficients, I can take the Inverse Fourier transformation to reconstruct the boundary with which I have started for which i have taken the forward transformation.

(Refer Slide Time: 19:53)




So, if you do that you find that if we had started with say a square shape something like this. Suppose, this square shape had capital N number of boundary points and from this capital N number of boundary points, we have also obtained capital number of Fourier coefficients then what I have done is, this capital N number of Fourier coefficients out of this I have considered say first P number of coefficients and if you know for the Fourier transformation you know that the low order coefficients Fourier coefficients gives you some information about the trend of the signal whereas, high order Fourier coefficients gives you the detailed information of the signal.

So, it may so happen that if I consider say P equal to 2 that is I consider only first two Fourier coefficients a u where u is equal to 0 and 1. So, only with these coefficients I want to reconstruct or find out the Inverse Fourier transformation for reconstruction of the boundary. So, for that my Inverse Fourier transformation expression will be sk is equal to a u e to the power j 2 pi u k upon capital N sum of this over u is equal to 0 to capital P minus 1 and k will vary from 0 to capital N minus 1.

So, what we are doing is, since I had capital N number of points on the boundary I am reconstructing capital N number of points to this Inverse Fourier transformation, but while doing. So, the number of Fourier coefficients that I am considering is not N number of Fourier coefficients rather P number of Fourier coefficients where P is lesser than N.

So, given this if I consider P equal 2 that is if I consider only the Fourier coefficients is u 0 and u 1 for reconstruction or in the Inverse Fourier transformation then maybe the kind of shape that I will reconstruct will be something like this which will be a circular shape because as I said that lower that coefficients gives you on the trend of the signal, but the high order coefficients gives you the details of the signal.

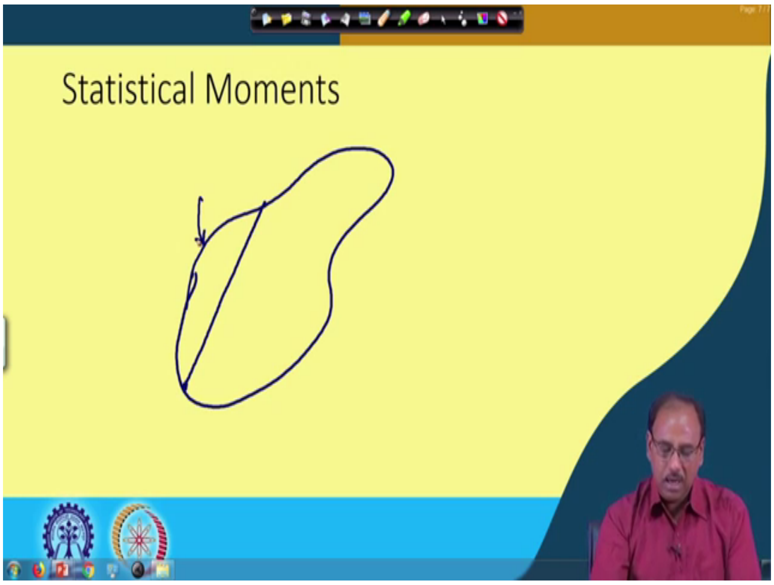
So, in a square shape the details means the presence of all these corners of the vertices. So, as I have truncated all the higher order coefficients, this detailed information in the reconstructed signal is lost. If I increase the value of P say from 2 if I go to 10 where I assume say P N is equal to something like say 128. So, against 128 if I take only 10 coefficients, I will get slightly better reconstruction.

So, which is like this it is not a perfect square neither a perfect circle. So, I get some detailed information present very constructed signal only when I consider all these 128 coefficients in the reconstruction or P varies u varies from 0 to 127 in this Inverse Fourier transformation expression then only I will get back my original shape which is a square. So, all these Fourier coefficients can also give you some information of or important information of the boundary of shape of the boundary.

So, over here I have obtained the different types of descriptors that I can get. Firstly, going for a polygonal representation when I go for polygonal representation then from the polygon itself I can compute some descriptors or when I go for polygonal represent representation then you find that the different vertices of the polygon that I get those vertices break the original boundary into a number of sub boundaries.

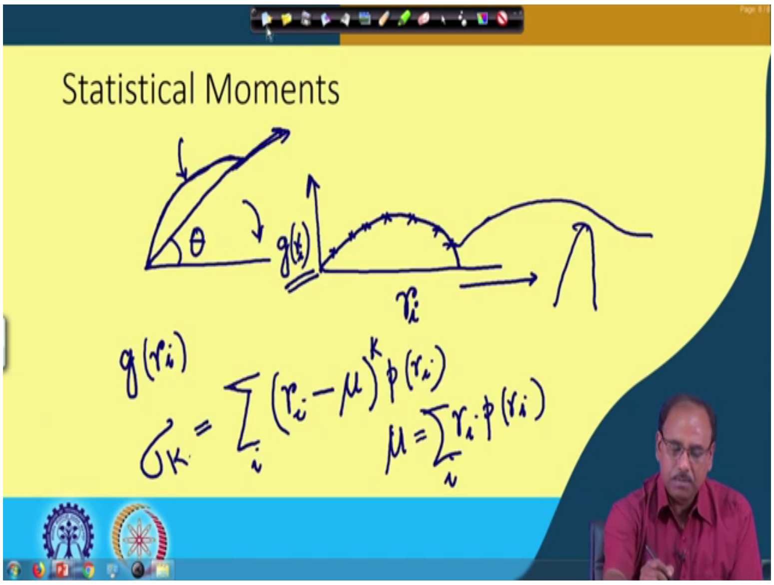
So, I can also obtain some shape information of all those sub boundaries then I can have what we have talked about is the signature and the nature of the signature or the shape of the signature gives you important information of the boundary. Then we have also talked about the Fourier descriptor which also gives you the important information of the shape of the boundary.

(Refer Slide Time: 25:30)




Now, let me go for some further analysis of the shape which is through statistical moments. So, what we have said is that say for example, I had an arbitrary shape something of this form and I had a polygonal representation through a polygonal representation, a segment of the shape which is this ok. So, one is I have this edge of the polygon, I also have this segment of the boundary. So, I can compute the shape of this boundary I can have different information about the shape of the boundary through statistical moments.

(Refer Slide Time: 26:16)




So, how I can do it? If I just put this shape like this. So, I have a shape of this form right this is a sub or segment of the arbitrary shape.

What I can do is this straight line I rotate it. So, that I rotate it by an angle theta in the clockwise direction. So, that this straight line chord or which is the edge one of the edges of the polygon that becomes horizontal. So, if I do that I will have boundary something like this. So, this is a boundary segment given this I can represent it in the form of a function.

So, I can put this as r and this function can be represented as g r. So, what I am doing is, I am representing this boundary or the boundary segment as a function g r; obviously, this will not be a continuous function because as I said before that I actually have a set of discrete points on this boundary segment. So, if I have if I represent this variable r by set of discrete variables say r i this boundary is represented by a function g r i.

Now, if I normalize this boundary normalize it by the area under this segment in that case this g r i is nothing, but a histogram. So, what I can say is that after normalization this g r i tells me what is the frequency of occurrence of this discrete variable r i in this in this particular scenario. So, it is a frequency of occurrence. So, once I have this frequency of occurrence then I can compute different types of statistical moments.

So, a statistical moment of order k is given by this r i is a minus mu to the power k times p r i where p r i is the frequency of occurrence or probability of occurrence of i take the summation of this over all i. So, this is a statistical moment of order k. So, I can put it as mu k sorry I will not put it as mu k. So, this one what I will do is instead of mu k as I have used mu to represent the mean, let me call it say sigma k.

So, sigma k is the statistical moment of order k given this particular distribution and you find that for different values of k, I can get a different types of shape information; obviously, in this case mu being the mean mu is nothing, but r i times p r i take the sum of this over all i. So, here you find that if I take the value of k is equal to 2 right. So, what I get is the variance.

If I take value of k to be equal to 3, I get third order moment. Second order moment is popularly known as the variance that tells you what is the spread of this distribution if the second order moment is very high; that means, the distribution will be of this form, it will be spread higher if the value of sigma square is small my distribution will be something like this. The third order moment tells you about the skewness or the symmetry of the distribution about its mean.

Similarly, all moments of different orders captures some information of the shape of this distribution or in this particular case, it captures the shape of this sub segment of this particular boundary. So, till now what we have covered is the different types of boundary information that we can obtain from a given shape. So, with this I conclude this part of the lecture. In the next lecture, I will talk about how we can obtain the different region descriptors or the region information including intensity, color and texture.

Thank you.