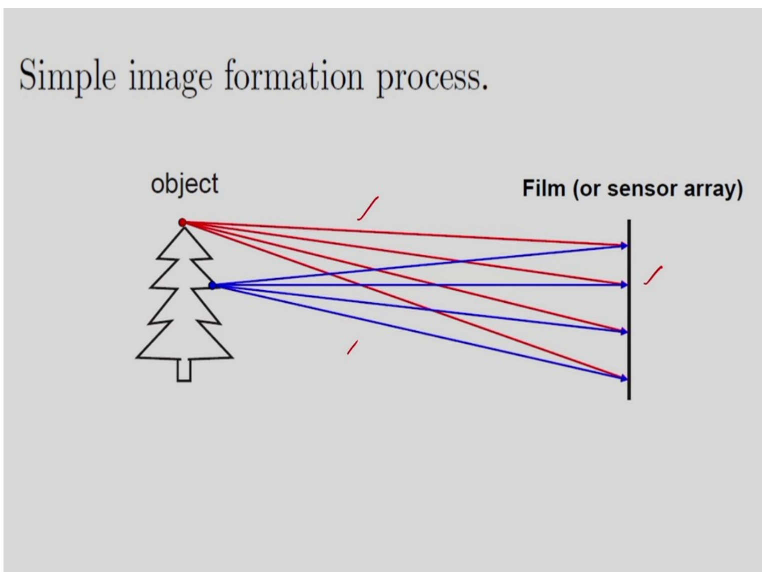
**Computer Vision and Image Processing – Fundamentals and Applications**
**Professor. Doctor M. K. Bhuyan**
**Department of Electronics and Electrical Engineering**
**Indian Institute of Technology, Guwahati**
**Lecture No. 06**
**Image Formation: Geometric Camera Models**

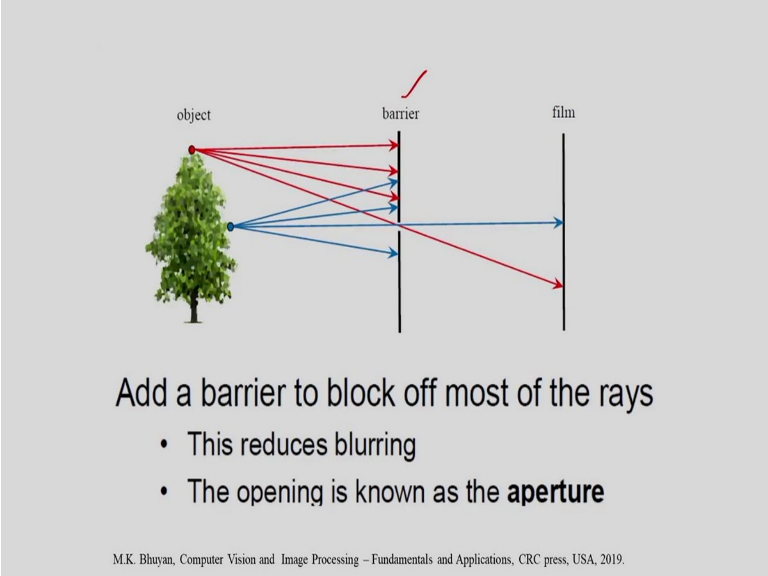
Welcome to NPTEL MOOC course on Computer Vision and Image Processing Fundamentals and Applications. So, in my last class I discussed about concept of affine transformation. I discussed some operations like translation operation, rotation operation, scaling operation. So, understanding of these operations is quite important to understand the concept of image formation in a camera. So, today I am going to discuss some important projection techniques in the camera. These are mainly the perspective projections, width perspective projections and also the orthographic projections. So, in my last class I discussed about the image formation in a setup.

(Refer Slide Time: 01:14)



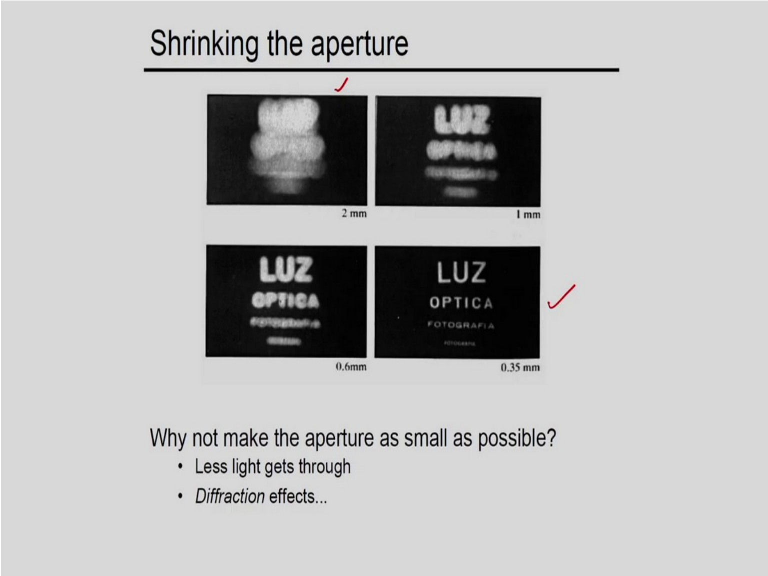

So, I have considered this set up. So, I have considered the object and this is the film. So, in this case I am getting the blurred image because of the intersections of the rays. If you, see the different rays are coming from the objects. And in this case I am getting the blurred image.

(Refer Slide Time: 1:35)




If I consider the another configuration, in this configuration I am considering one barrier. I have the object here. I have the sensor area here. So, if I consider this aperture, aperture is nothing but the small opening. Then in this case I am getting the good quality image. This aperture actually reduces blurring. This small opening is called the aperture. In my previous slide I have shown the case. In this case, in that case I was getting the blurred image. But in this case I am getting at least some good quality image because of the aperture.

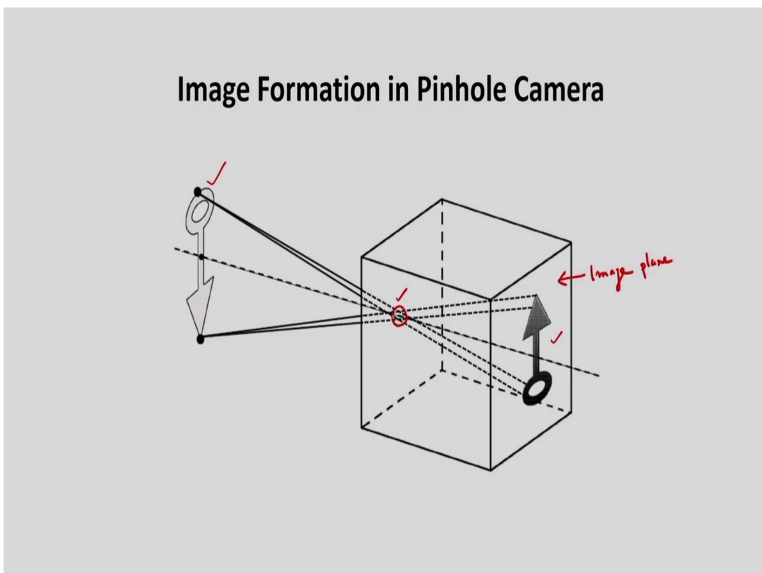
(Refer Slide Time: 2:05)



I have shown some examples like, I am considering the aperture size is 2 mm, aperture size is 1mm, aperture size is 0.6 mm, aperture size is 0.35 mm. If I reduce the aperture then what is happening? I am getting the sharp image like this. This is the blurred image corresponding to the aperture size of 2 mm. And this is the image I am getting the 0.35 mm, the sharp image I am getting.

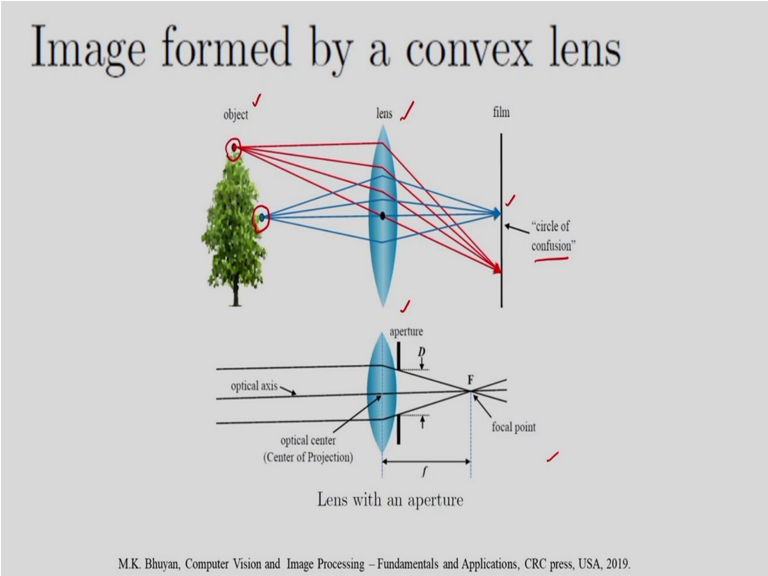
But I cannot make a aperture as small as possible. Why? Because less light it can pass through the aperture. And also I have to consider the effect of diffraction. So, that is why I cannot make the aperture as small as possible.

(Refer Slide Time: 02:46)




And here I have shown the image formation in a pinhole camera. So, this is a very simple camera configuration. This is my object. I am getting the inverted image in my image plane. This is my image plane and this is one small opening. So, this configuration is very similar to our human eye. So, this is nothing but the pupil of the eye. And this image plane is nothing but retina of the eye. So, we are getting the inverted image. The human brain reconstruct the 3D information. Because it is 3D to 2D projection so human brain reconstruct the 3D information.

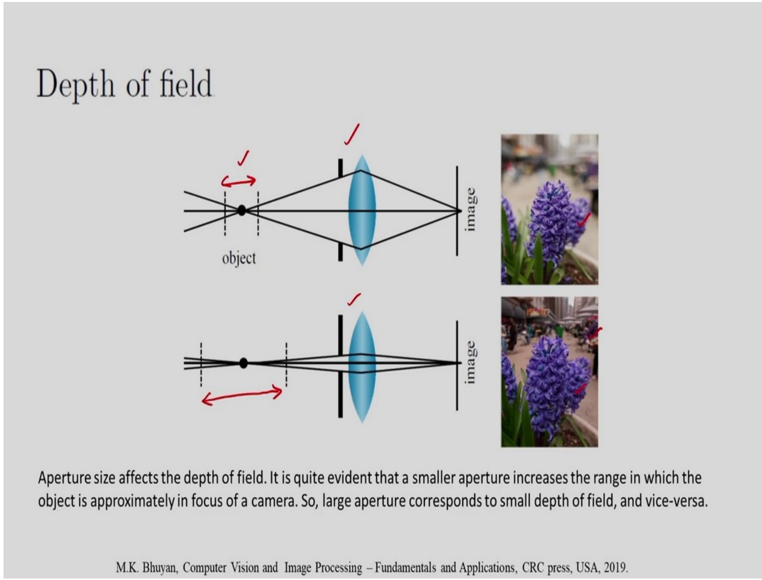
(Refer Slide Time: 03:28)




Last class also I had shown the image formed by convex lens. So, here I have shown the convex lens. This is my convex lens. And I have the object. And in this case if you see, this portion of the object is properly focused. So, I am getting the image, the sharp image. But if I consider at this portion of the object, the second, this portion, this is not properly focused. So, that is why corresponding to that portion I am getting the circle of confusion. I am getting the blurred image corresponding to that portion.

In the second diagram, if you see this diagram, I am considering the convex lens with an aperture. So, I am considering the aperture and I have shown the focal length. f is the focal length and the focal point also I have shown and this is the configuration to get the image. Optical axis also I have shown. So, lens with an aperture.

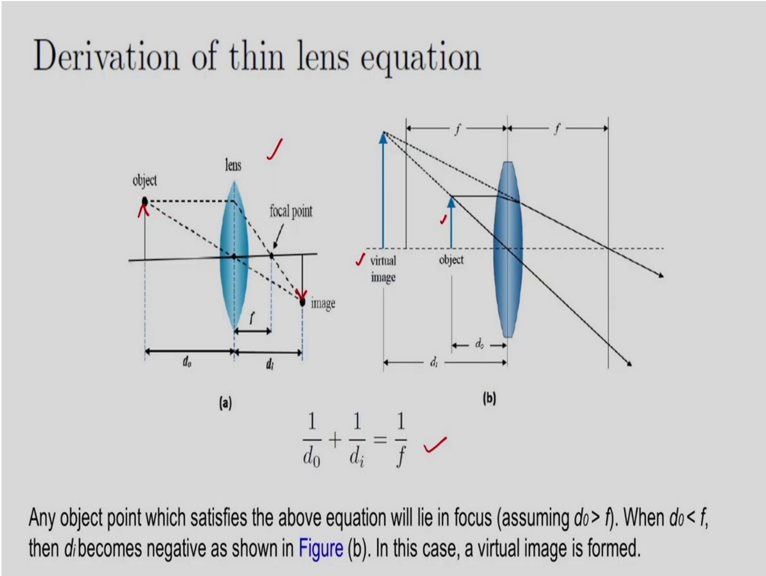
(Refer Slide Time: 04:24)




Next, in the last class I discussed about the concept of the depth of field. The depth of field depends on the aperture size. So, in the first case I had considered the wide aperture. In the second case I am considering the small aperture. Corresponding to the wide aperture only this portion, if I consider this range, from this to this range, this range is properly focused. Corresponding to the small aperture if you see the range, so this range, this range is properly focused that means the aperture size affects the depth of the field. So, a small aperture increases the range in which the object is approximately in focus of a camera. So, large aperture corresponds to small depth of field and vice versa.

So, in my first case, only the, this range I am considering. That means I am getting the foreground in the image. The background is the blurred because that is not focused properly. In the second case the range is high. So, I am getting the foreground and the background. So, you can understand the depth of field depends on the aperture.

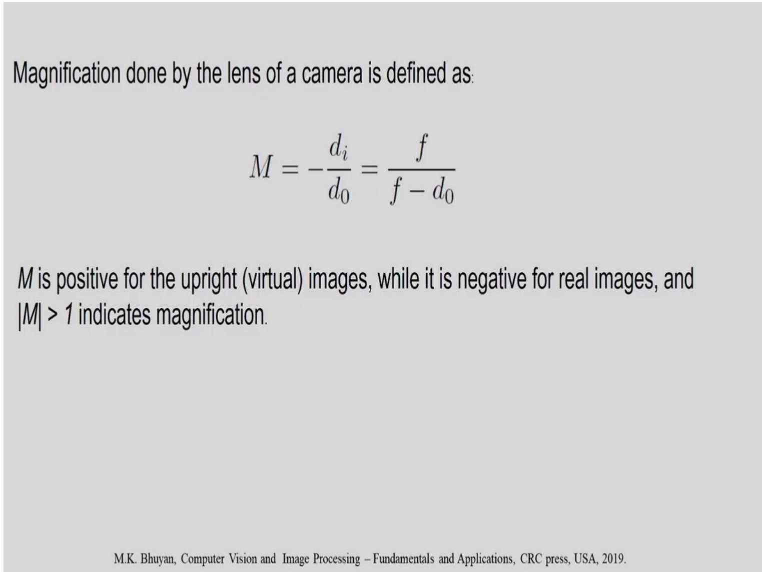
(Refer Slide Time: 05:27)




And last class also I have show the derivation of thin lens equation. So, in the first case I have shown the real image and in this case this is the object and I am getting the inverted image. Here I am shown the focal length and the distance between the image and the lens that is di and distance between the object and the lens that is $d_0$I have shown. In the second case I have shown the virtual image. So, virtual image is this. I have shown the object, object is this. So, from here if you see this, the configuration I can get the, the lens equation. The lens equation is 1 by $d_0$ + 1 by di is equal to 1 by f. So, already you know this equation.

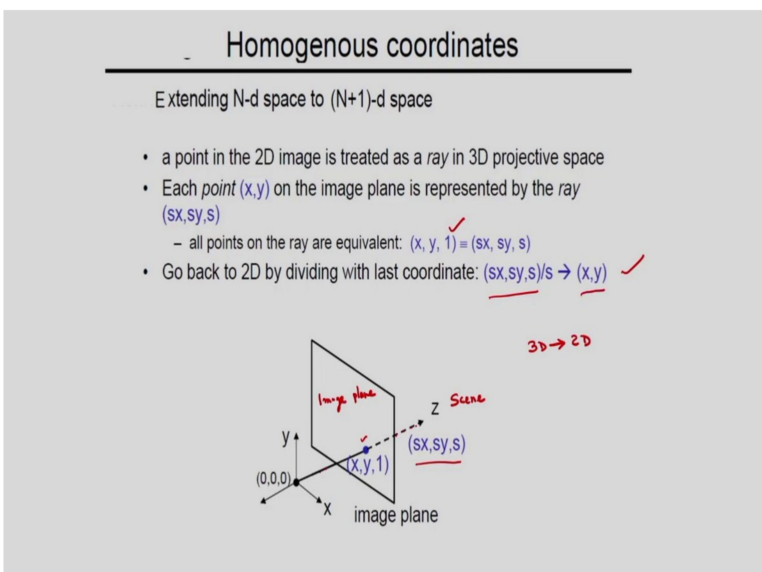
Now in the second case if you see, $d_0$< f. Then in this case di will be negative. And in this case I am getting the virtual image. So, in the second case I am getting the virtual image. In the first case I am getting the real image but that is inverted image.

(Refer Slide Time: 6:23)




And also I have defined the magnification done by the lens of a camera. So, magnification is defined by M is equal to minus di divided by $d_0$. So, that is the magnification. So, magnification is considered as positive for the virtual image. And for the real image it is negative and if I consider M is greater than 1 that corresponds to magnification.

(Refer Slide Time: 06:47)

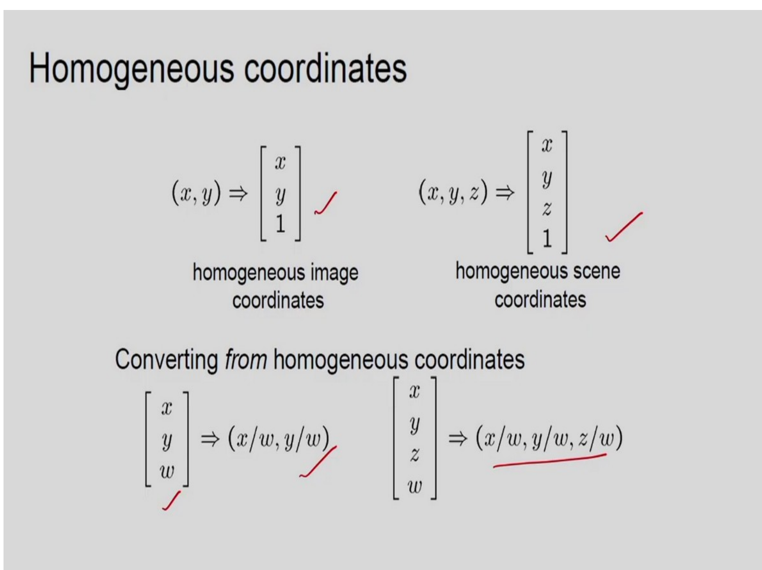



Now let us consider homogenous coordinate. I discussed this homogenous coordinate in my affine transformation class, the geometric transformation. So, homogenous coordinate means extending n-dimensional space into n plus 1 dimensional space. So, in this diagram I have shown, one is the image plane. This is my image plane. And I have shown this is the scene, this is the scene. So, a point in the 2D image is treated as a ray in 3D perspective space. So, point in the image plane, the point is x comma y. This is the point in the image plane that is considered as a ray in 3D perspective space, so, what is the 3D perspective space? So, I am considering as a ray sx comma sy comma s. So, this is the, in the perspective space.

So, that means if I consider all the points in the ray so, if I consider all the points in the ray that corresponds to the point x comma y comma 1 in the image plane. So, that means it is nothing but the 3D to 2D projection, this is nothing but the 3D to 2D projection because corresponding to all the points in the ray I am getting only one point in the image, that point is x comma y comma 1.

And in this case if I have the ray coordinate, suppose this is the, in the perspective space if I want to get the image coordinate what I have to do? I have to divide sx by s, that is by the third coordinate, sy divided by s and s divided by s. And after this I have to neglect the third coordinate. Then in this case I am getting the x comma y. So, I can get the image coordinate from the scene coordinate like this.

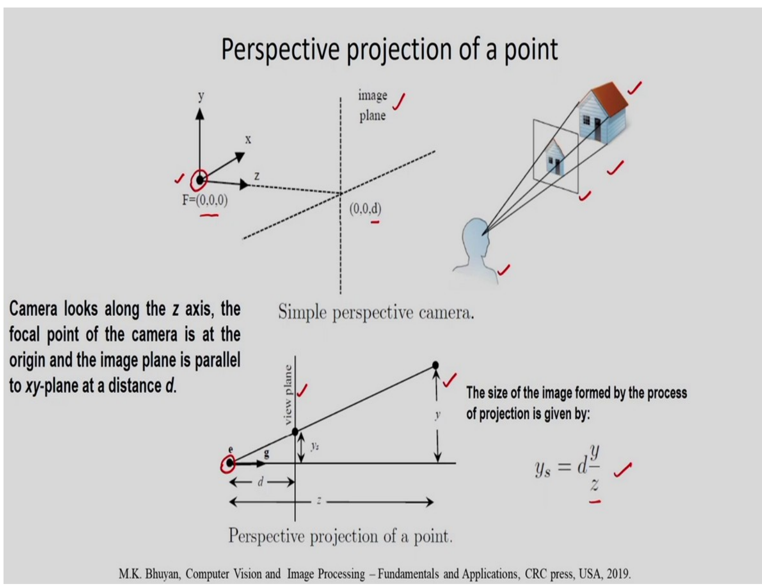
(Refer Slide Time: 8:39)




The x comma y I can write in homogenous coordinate like this, x y 1, adding one extra coordinate and if I consider it the scene coordinate x y z, that can be represented in the homogenous coordinate x y z 1. So, I can represent like this. And suppose if I want to convert from homogenous coordinate how to do the conversion? So, this is in the homogenous coordinate. If I want to do the conversion from homogenous to the image coordinate then in this case what I have to do?

The x coordinate is divided by the third coordinate. The third coordinate is w. y coordinate is divided by the third coordinate and third is also divided by would and after this I am neglecting the third coordinate. So, I am getting the coordinate x divided by w, y divided by w. And similarly in this case also I am considering like this x divided by w, y divided by w, z divided by w. And I am neglecting the fourth coordinate, the fourth coordinate is 1. So, I am neglecting this 1.

(Refer Slide Time: 09:36)



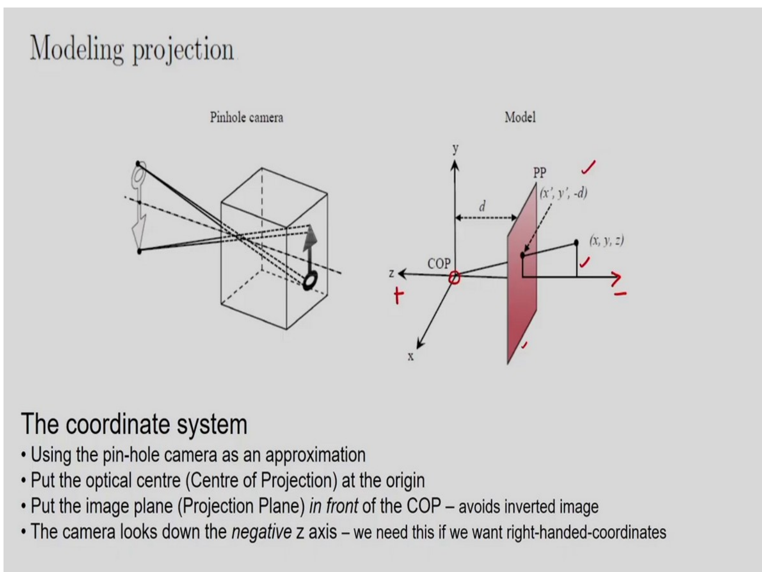

So, let us consider the case of perspective projection of a point. So, here I have shown, in the first diagram if you see I have shown the camera, this is the camera centre of projection. Camera looks along the z axis. The focal point of the camera is at the origin, that is origin is 0 0 0. So, I am considering this one that is the focal point. And image plane is parallel to xy plane. So, I have this image plane that is parallel to xy plane.

And what is d? The d is the distance between the centre of projection and the image plane. So, this is my centre of projection that is the focal point of the camera. And here I have shown one example of the perspective projections. So, my object is this, this is my image plane and this is my camera, camera means the viewing direction.

And in the third case I have shown the perspective projection of a point. So, first I have shown the centre of projection, this is the centre of projection. And I have considered the object, object is y. And this is the view plane. View plane is nothing but the image plane. So, what is the distance between the centre of projection and view plane? The distance is d. The distance between centre of projection and the object is z.

So, from this you can see I can get this expression ys is equal to dy divided by z. So, what is ys? ys is nothing but the size of the image in the image plane; that is the perspective projection of a point. If I increase z, what is z?  The distance between the centre of projection and the object. If suppose it is some high value then what will be the image size? The image size will be reduced. That is the concept of the perspective projection.

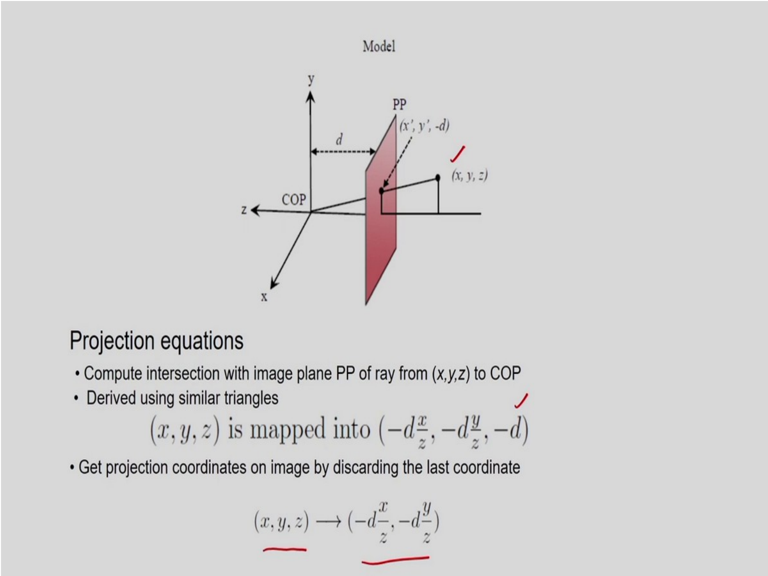
(Refer Slide Time: 11:18)




So, here I am shown again the modeling of projection. Again I have shown the pinhole camera model as an approximation. And after this I have shown the projection model. So, in this case I have considered the optical centre that is the centre of projection at the origin. So, COP means the centre of projection and I have considered the object. The object is x y z. And this PP means the plane of projection that is the image plane.

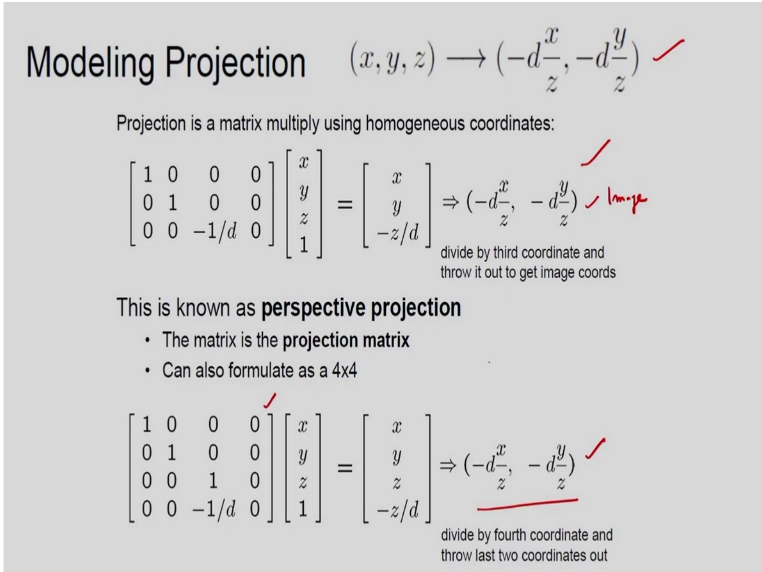
Put the image plane that is the projection of plane in front of the COP. Why I am putting this one? Because I want to avoid inverted image. That is why I am putting the plane of projection in front of the COP. The camera looks down the negative z axis. And this is the positive z axis. So, camera looks down the negative z axis. So, what is the size of the image in the PP, the plane of projection? The size is x dash y dash and minus d. Why it is minus d? Because I am considering negative z direction.

(Refer Slide Time: 12:18)



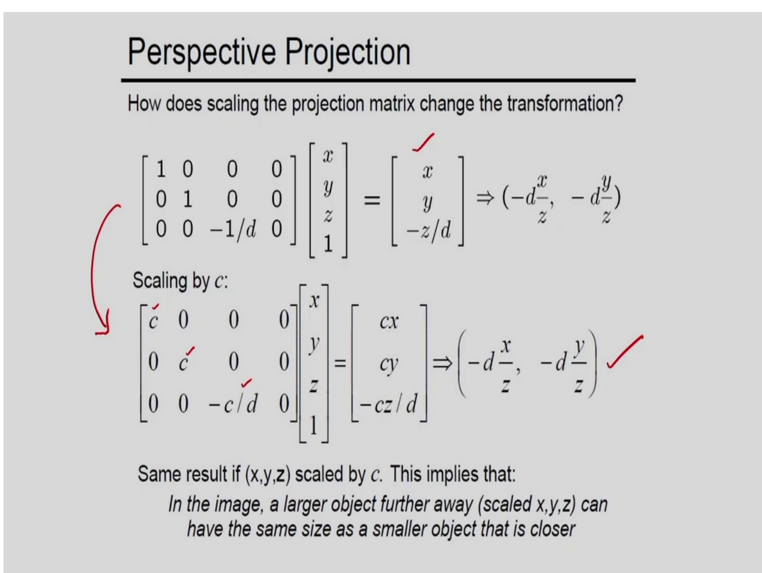

Now if I consider same model, so if I consider the projection equations. So, if I see the similar triangle, I have two similar triangles here if you see, one is the big triangle, another one is a small triangle. So, if I consider these triangles you can see the point x y z that is the scene coordinate is mapped into the image plane, the image point will be minus d x divided by z minus d y divided by z minus d. I have to neglect the third coordinate. So, that means the point x y z is mapped into minus d x divided by z minus d y divided by z in the image plane. So, this is the concept of the perspective projection. So, how to define this equation in the matrix form?

(Refer Slide Time: 13:03)




So, in my next slide I should this equation. This is the equation corresponding to the perspective projection. This, the mapping I can represent in the matrix form, the matrix form is this. So, this is my image coordinate, image coordinate I am getting. This is called the perspective projection matrix, the projection matrix. This matrix can be formulated as a 4 by 4 matrix. So, if you see here I am considering the 4 by 4 matrix because the previous matrix equation, that is not symmetric expression. But in this case if I consider the 4 by 4 formulation that is the symmetric matrix. So, that is my perspective projection.

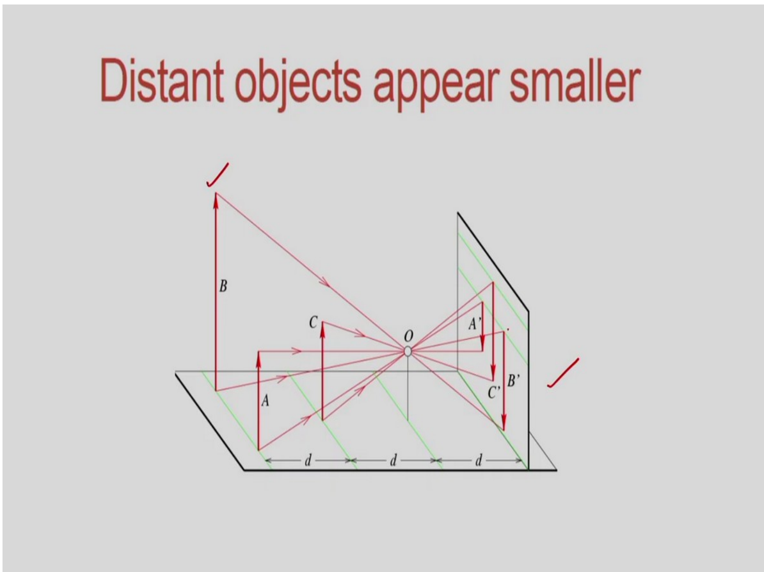
(Refer Slide Time: 13:42)



Now let us consider how does scaling the projection matrix change the transformation? So, this is my perspective projection. Now let us consider scaling of x coordinate by c, y coordinate by c and z coordinate by c from the previous equation. Corresponding to this I am getting the same output. What is the output? The output means the image coordinate. The image coordinate will be minus dx divided by z minus dy divided by z.

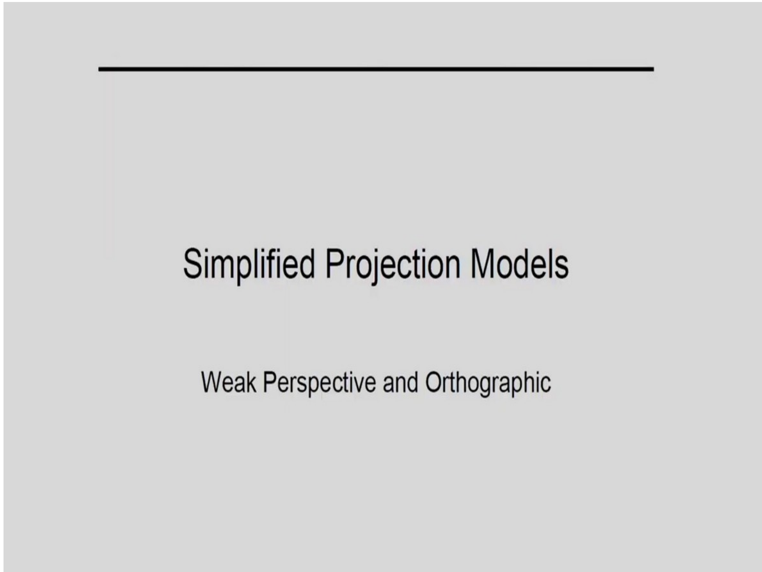
That means, what is the meaning of this? In the image, a larger object further away, that is scaled by x comma y comma z, that x coordinate is scaled by c, y coordinate is scaled by c, z coordinate is scaled by c, can have the same size as a smaller object that is closer. So, that is the interpretation of this case, scaling by c in a process perspective projection matrix.

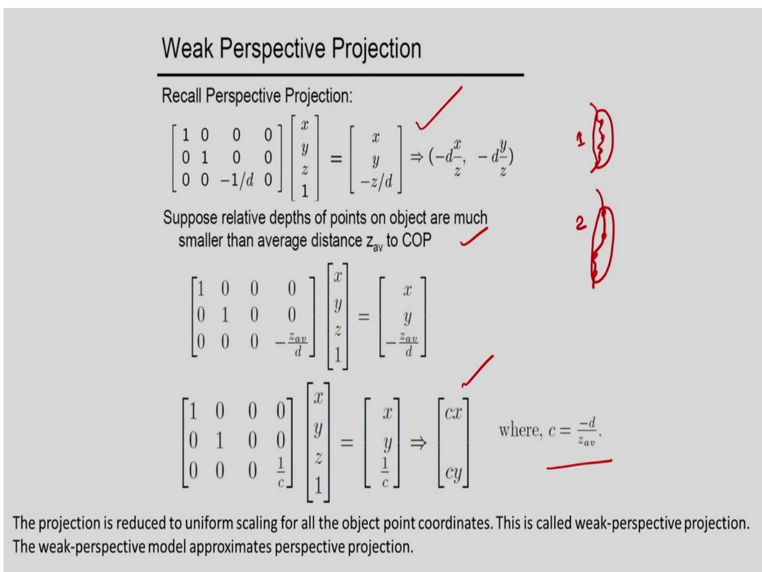
(Refer Slide Time: 14:33)




So meaning is; the physical interpretation is the distant objects appear smaller. So this is the, the outcome of the perspective projection. So, if I consider this object the distant object appear smaller in the image plane. This is my image plane.

(Refer Slide Time: 14:50)



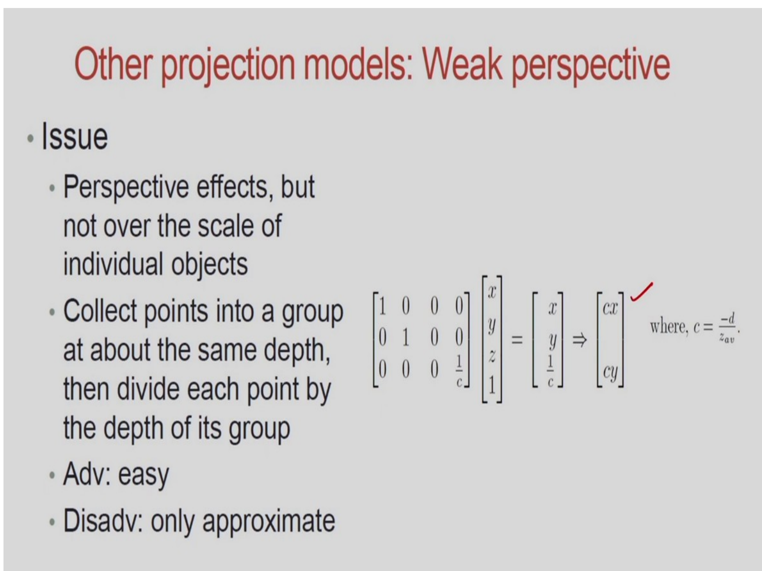




After this I will discuss some simplified projection models. That is called the weak perspective projection. Another one is the orthographic projection. So, what is weak perspective? Here you see. So, we have this matrix formulation for the perspective projection. And suppose if I consider a surface. Suppose one surface is like this. So, suppose these points if I consider, the relative depth of these points are much smaller than the average distance z av to COP. So, suppose relative depth between these points are much smaller, then these points I can considered as a group.

Suppose I have another object like this. So, this is my object number 1 suppose in the image. And this is my another object, the object number 2. So, if I consider these points, the relative depth of the points on the objects are much smaller than the average distance, then in this case I can considered as a group, I can consider as a same point. Now if I considered this case, if you see, if this consider, if I consider this case, suppose relative depth of points on objects are much smaller than the average distance z av to COP, centre of projection then I have this equation.

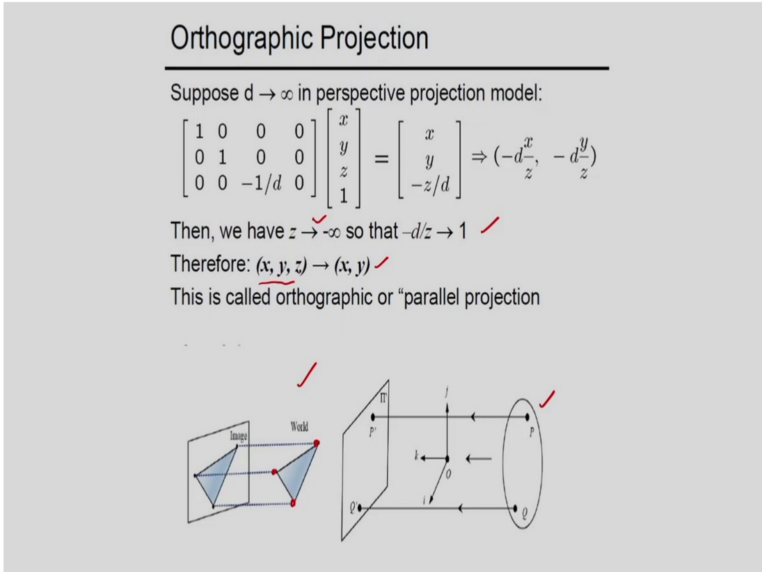
And if I consider c is equal to this one, c is equal to minus d divided by z av then in this case I have this one. That means x-coordinate is scaled by c. And y-coordinate is scaled by c. This is nothing but simple scaling. The projection is reduced to uniform scaling for all the object points coordinate and this concept is called the weak perspective projection. So it actually, the weak perspective model approximate the perspective projection model. So it is a scaling of x coordinate and scaling of y coordinate.

(Refer Slide Time: 16:41)




So, the meaning of this, it will give the perspective effect but not over the scale of individual objects. And in this case I had to collect the points into a group that I have shown in previous slide. So, collect the point into group at about same depth and then divide each point by the depth of that group. So, it is the same thing, the scaling of x coordinate and scaling of y coordinate. And this concept, the weak perspective projection approximate the perspective projection.

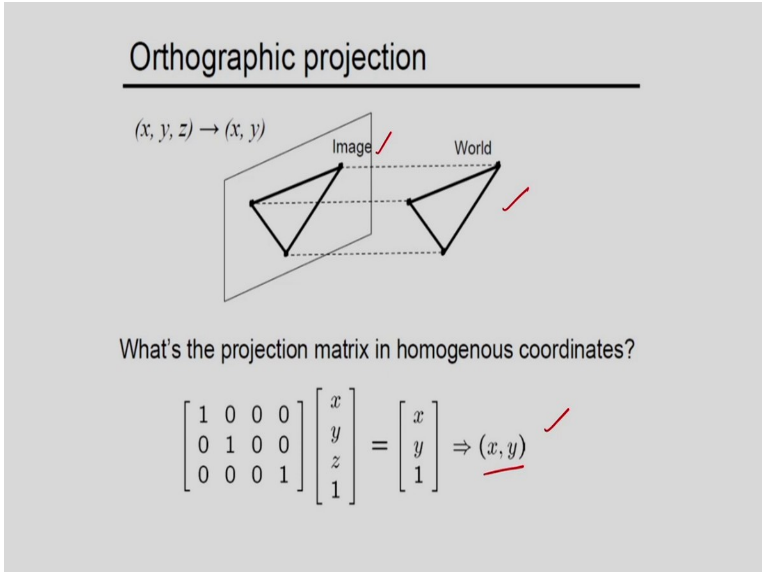
(Refer Slide Time: 17:10)




And the next one I have considered the orthographic projection. In orthographic projection if I consider d tends to infinity, the d is the distance between the centre of projection and the image plane. It is very high. And also I am considering z tends to minus infinity, minus I am considering because I am considering the negative z direction. If z is very high, z is the distance between the centre of projection and object. That is also very high, then in this case minus d by z tends to 1.

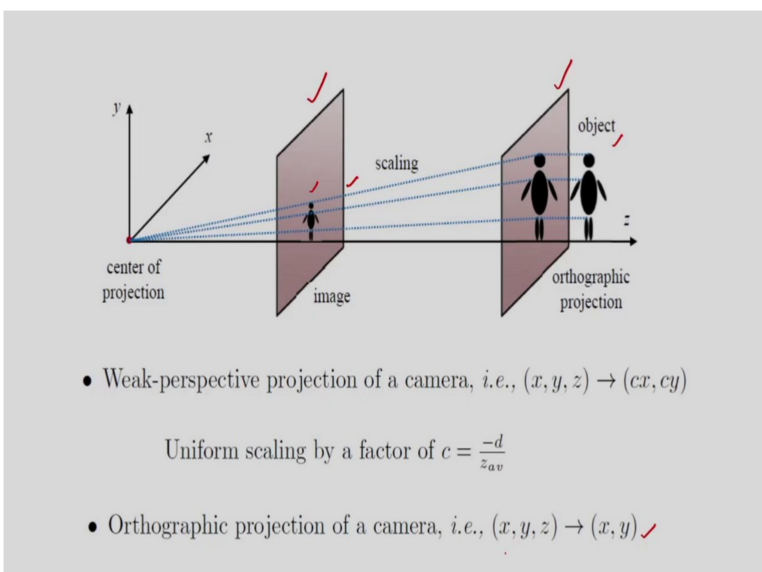
And if I consider this case if minus d by z tends to 1, that means scene coordinate is x comma y comma z that is mapped into x comma y in the image plane. This is nothing but simply 3D to 2D projection. The coordinate x y z is mapped into x comma y. This is called the orthographic projection or the parallel projection. So, here in this figure I have shown the old coordinates, scene coordinates are like this. The coordinates is x y z, but in the image I have the two coordinates, only x comma y. Again I am showing the same concept here. The point x y z is mapped into x comma y.

(Refer Slide Time: 18:18)




In this diagram also I have shown the same concept. The point x y z is mapped into x comma y. That is the, from world coordinate to the image coordinate I am doing the mapping. And what is the projection matrix in homogenous coordinate? In homogeneous coordinate I can write the, the projection matrix in this form. That is the orthographic projection. So, this is my image coordinate.

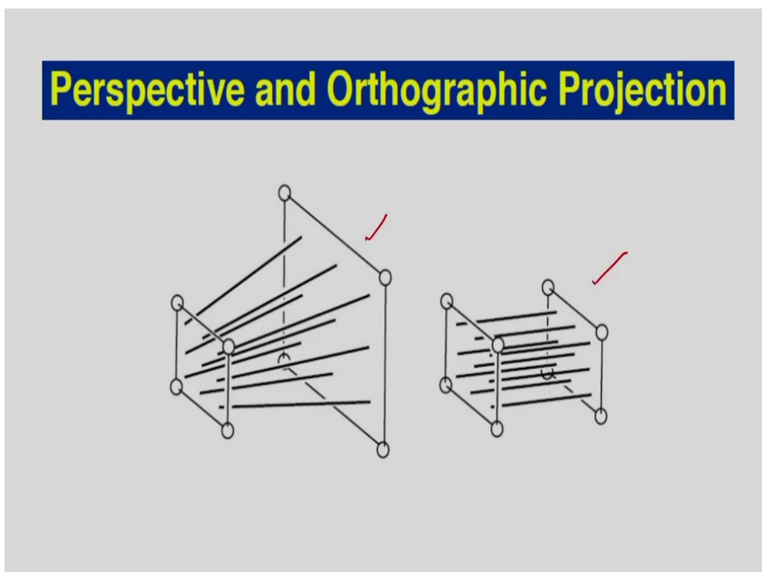
(Refer Slide Time: 18:40)

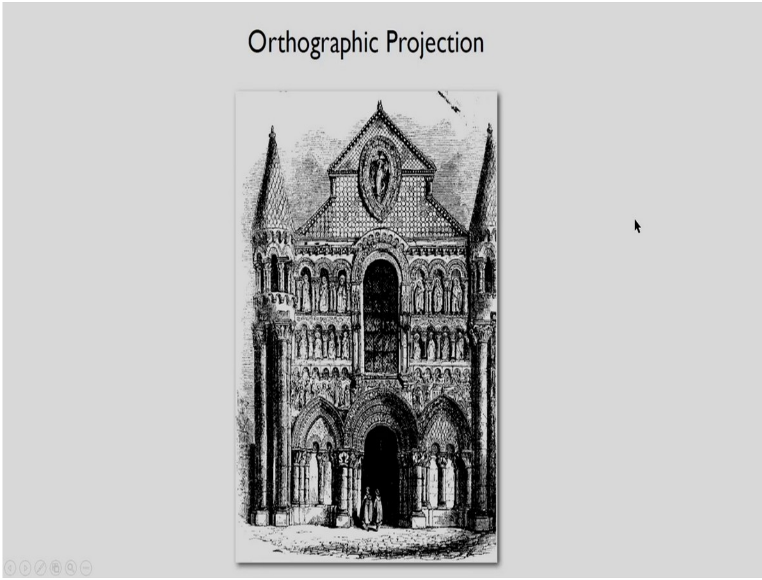


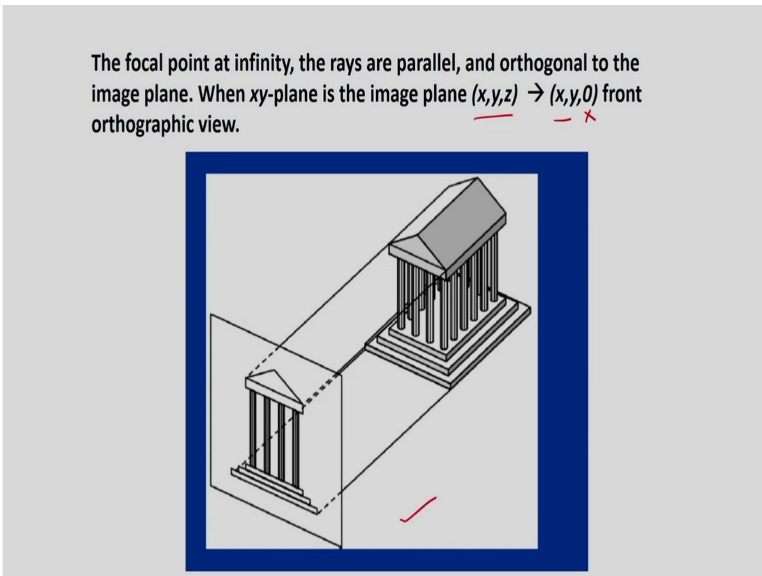

This concept of the orthographic projection and the weak perspective I have shown in this diagram. This is my image plane, suppose and this is my object. In case of the, which perspective projection what I am getting? I have the scaling of x coordinate and scaling of y coordinate. So, that is weak perspective projection and I am getting the image corresponding to weak perspective projection.

In case of the orthographic projection what is my consideration? The consideration is the distance between the image plane and the centre of projection is very high. And the distance between the centre of projection and the object is also very high. In this case the point x y z is mapped into just the point x comma y. So, in this case I am considering the image plane is here. The distance between the center of projection and the image plane is very high and distance between the centre of projection and the object is also very high. And in this case I have to consider the orthographic projection. So, this is my orthographic projection.

(Refer Slide Time: 19:32)






And here I have given two examples. One is the, first one is the perspective projection. And second one is the orthographic projection. And I am shown one image corresponding to orthographic projection. In this case there is no z information because it is mapping from x y z to x comma y.

And finally I want to show one example of the orthographic projection. The focal point at infinity, the rays are parallel and these are; rays are orthogonal to the image plane. Then in this case this is the mapping from x y z to x y. And in this case the z information is missing. We do not have the z information. And I am getting the orthographic view.

So, in this class I discussed the concept of projection. So, 3 projection techniques I have discussed. One is the perspective projection, one is the weak perspective projection, another one is the orthographic projection. So, these concepts are very important. So, in the next class I will discuss the concept of camera calibration. So, camera has extrinsic parameters and the intrinsic parameters. And based on this how to do the camera calibration?  So, that concept I am going to explain in the next class. So, let me stop here today. Thank you.