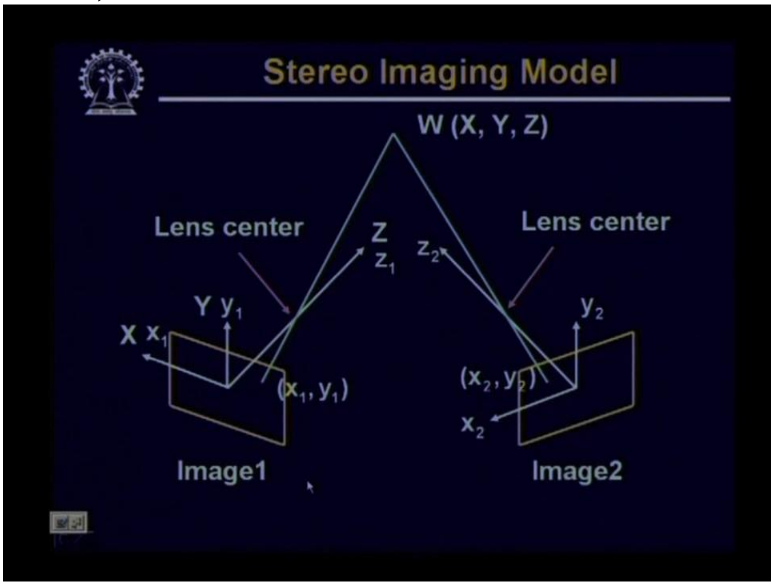
Digital Image Processing
Prof. P. K. Biswas
Department of Electronics and Electrical Communications Engineering
Indian Institute of Technology, Kharagpur
Module 03 Lecture Number 15
Stereo Imaging Model - 2

(Refer Slide Time 00:19)

Hello, welcome to the video lecture series on Digital Image Processing. So this camera calibration using this procedure can be done for any given imaging setup. But the problem still exists, given an imaging point, an image point I cannot uniquely identify what is the location of the 3D world point. So for the identification of the 3D world point or finding out all the 3 x, y and z coordinates of a 3D world point I can make use of another camera. So let us look at a setup like this

(Refer Slide Time 00:58)

(Refer Slide Time 00:59)

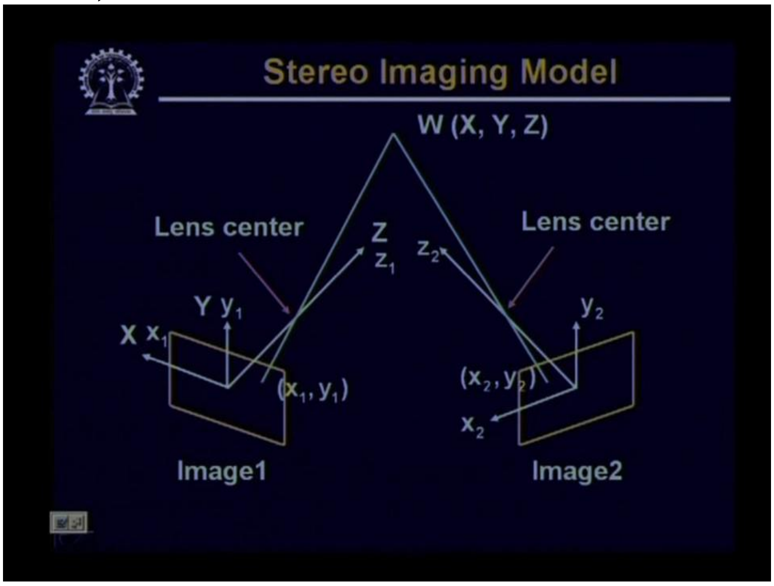


where on the left side I have image 1, on the right side I have image 2. The image 1 is taken, captured with the help of one camera; image 2 is taken with the help of another camera. So image 1 has the coordinate system say x 1 y 1 z 1. Image 2 has the coordinate system x 2 y 2 z 2. And we can assume that the 3D world coordinate system that is capital X capital Y capital Z is aligned

(Refer Slide Time 01:32)

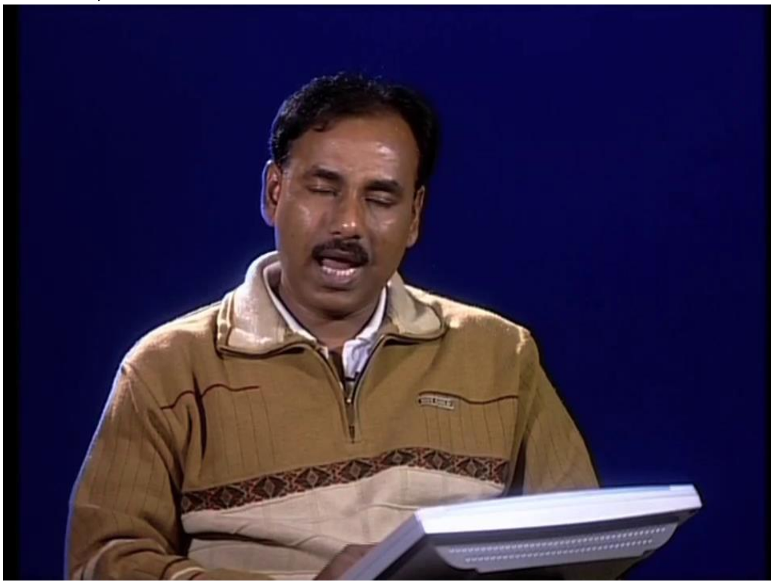
with the left camera that means the left image coordinate system is same as the 3D world coordinate system where as the right image coordinate system is different. Now once I have this, given a point

(Refer Slide Time 01:47)



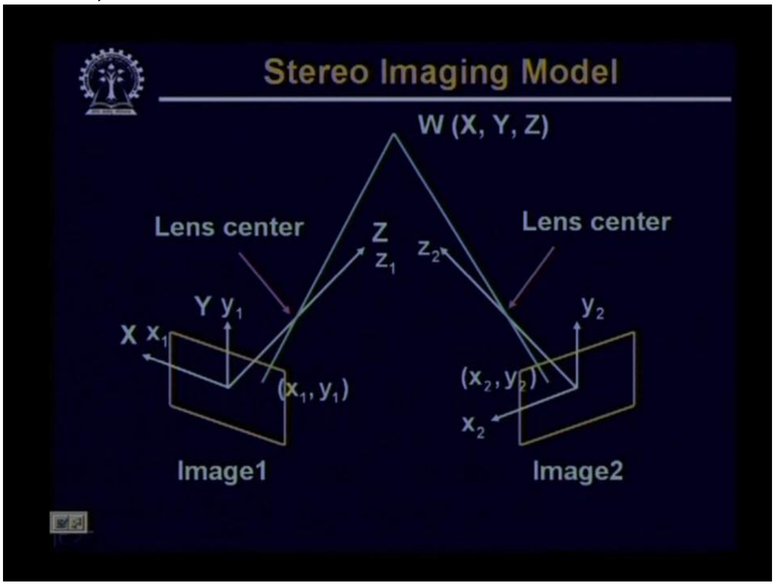
w in the three-dimensional world, three-dimensional space you find that the corresponding image point in image 1 is given by x 1 y 1 and the image point for the same point w in image 2 is given by x 2 y 2. I assume that both the cameras are identical that means they have the same value of wavelength lambda so they will have the same perspective transformation as well as

(Refer Slide Time 02:16)



inverse perspective transformation Now once I know that in image 1, the image point corresponding to point w is x 1 y 1, then by applying the inverse perspective transformation I can find out the equation of the straight line on which the three-dimension, the point w will exist. Similarly from image 2 where I know the location x 2 y 2 of the image point if I apply the

(Refer Slide Time 02:46)

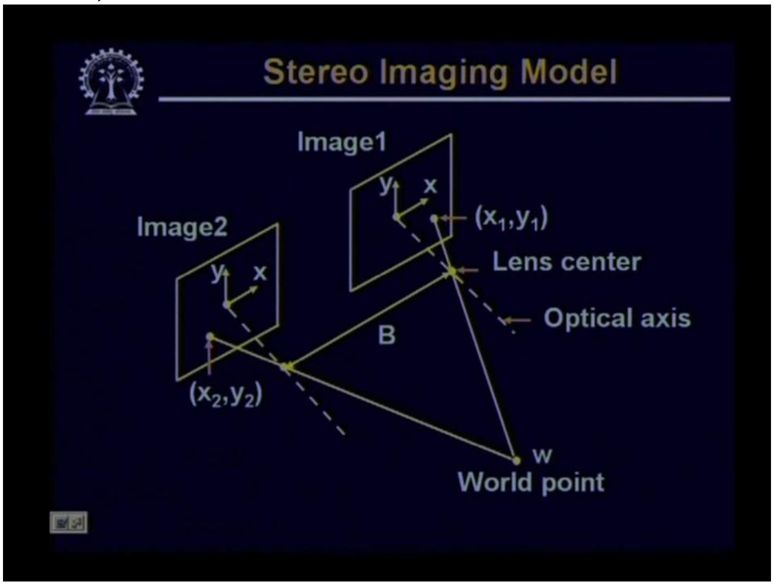


inverse perspective transformation I also get equation of another straight line on which this equation of w will exist. So now you find that by using these two images I got equations of two straight lines. So if I solve these two equations, then the point of intersection of these two straight lines gives me the x y z coordinate of

(Refer Slide Time 03:12)

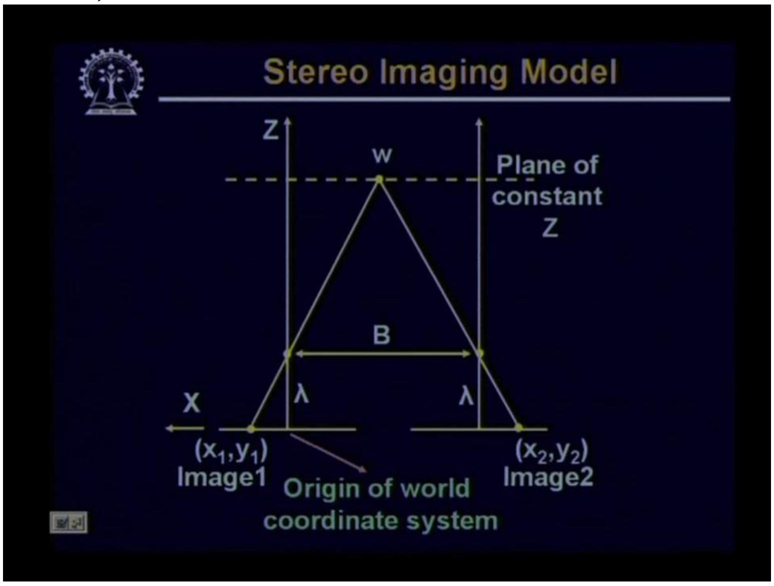
point w But here you find that we have taken a general stereo imaging setup where there is no alignment between the left camera and the right camera or between the first camera and the second camera. So for doing all the mathematical operations what we have to do is we have to apply again a set of transformations to one of the camera coordinate systems so that both the camera coordinate systems are aligned. So these transformations will again involve, may be a transformation for some translation, transformation for some rotation and possibly it will also employ some transformation for scaling if the image resolution of the, image resolution of both the cameras are not same. So there will be set of transformations and the corresponding mathematical operations to align the 2 camera systems .but here you find that positioning of the camera is in our control. So why do we consider such a generalized set up. Instead we can arrange the camera in such a way that we can put the imaging plane of both the cameras to be coplanar. And we use the coordinate systems in such a way that the x coordinate system and of one camera and the x coordinate system of the other camera are perfectly aligned. There will be a displacement and the displacement in the z axis. So effectively the camera setup that we will be having

(Refer Slide Time 04:57)



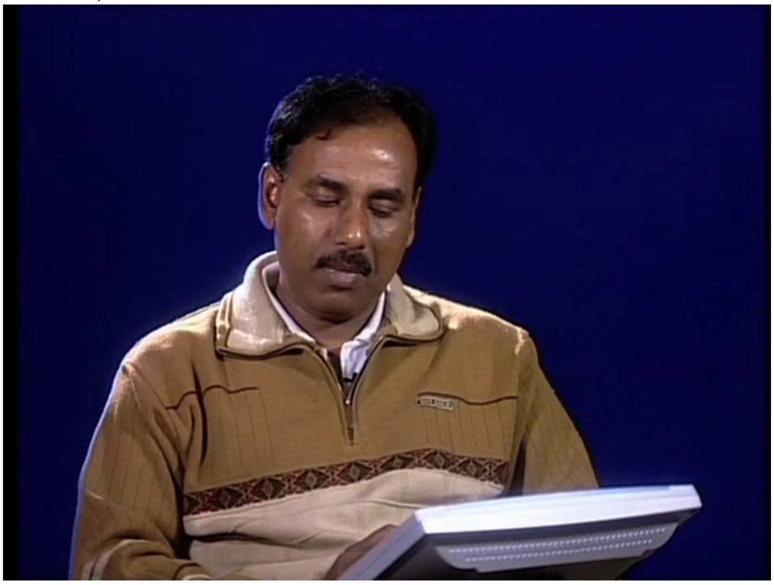
is something like this Here you find for the two cameras, the image plane 1 and image plane 2, they are in the same plane, the x axis of both the cameras the camera coordinate systems they are collinear. The y axis and z axis, they have a shift of value b. So this shift b, this value b is called the camera displacement. We assume both the cameras are identical otherwise, that is they have the same resolution, they have the same focal length w. Again here the 3D point w we have in image 1, the corresponding image point as x 1 y 1 and image 2 we have the corresponding image point as x 2 y 2. Now this imaging setup can be seen

(Refer Slide Time 05:55)



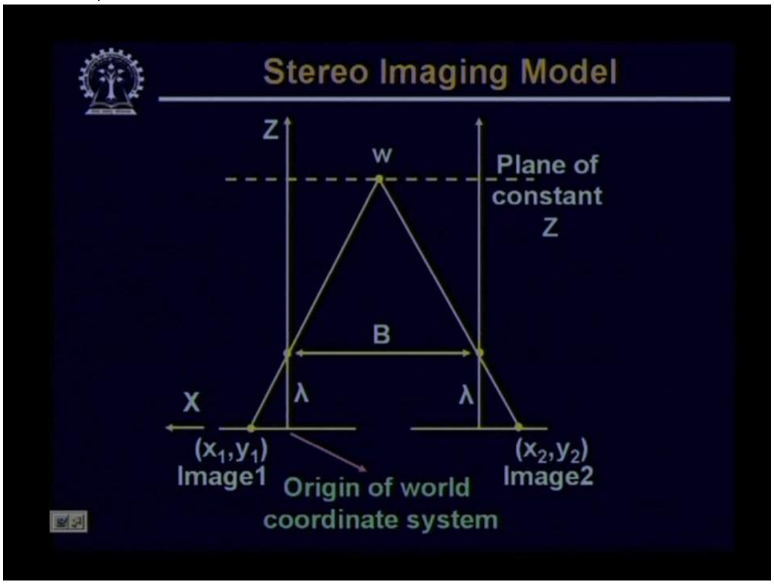
as a setup where find that this x y plane of both the cameras are now perpendicular to the plane So I have x axis which is horizontal, the z axis which is vertical and the y axis which is perpendicular to this plane So in this figure I assume that the camera coordinate system of one of the cameras, in this case the camera 1 which is also called the left camera is aligned with the 3D world coordinate system

(Refer Slide Time 06:31)



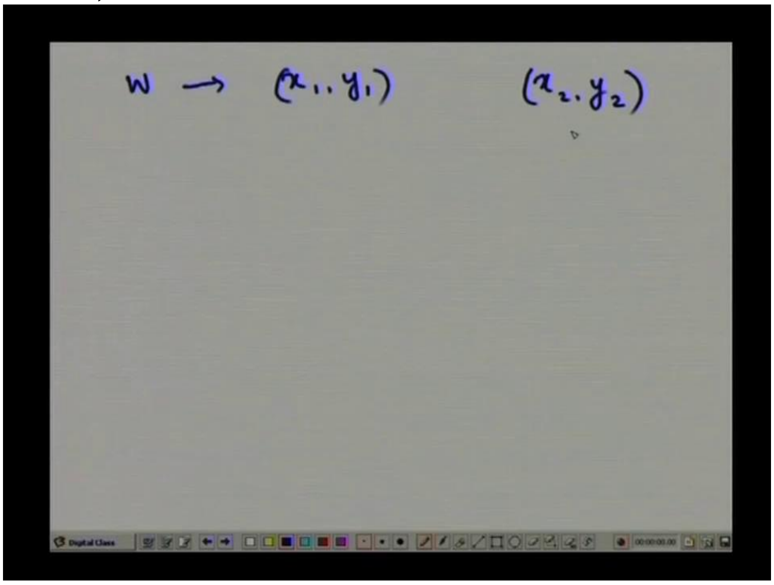
capital X capital Y capital Z, the coordinate system of the left camera is assumed to be x 1 y 1 z 1. The coordinate system of the right camera is assumed to be x 2 y 2 z 2. Now given this particular imaging setup you find that for any particular image point say w, with respect the cameras, camera 1 and camera 2, this point w will have the same value of the z (x) coordinate. It will have the same value of the y coordinate but it will have different values of the z coordinates because the cameras are shifted or displaced only in the z axis, not in the x axis or y axis. So origin of this

(Refer Slide Time 07:19)



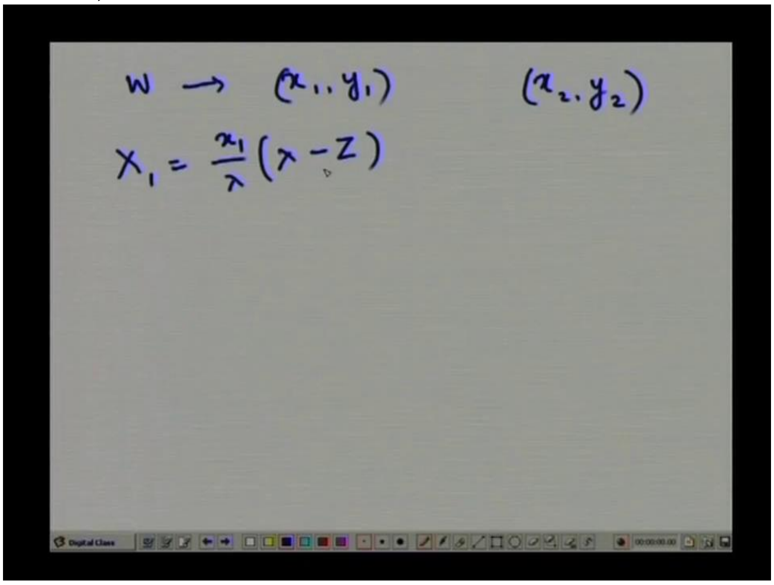
world coordinate system and origin of the left coordinate system, they are perfectly aligned. Now taking this particular imaging setup, now I can develop a set of equations. So the set of equations will be something like this. We have seen that for image 1, for point w the corresponding image point is at location x 1, y 1. For the same point w in right image, the image point is at location x 2 y 2. So these are the image coordinates in the left camera and the

(Refer Slide Time 08:03)



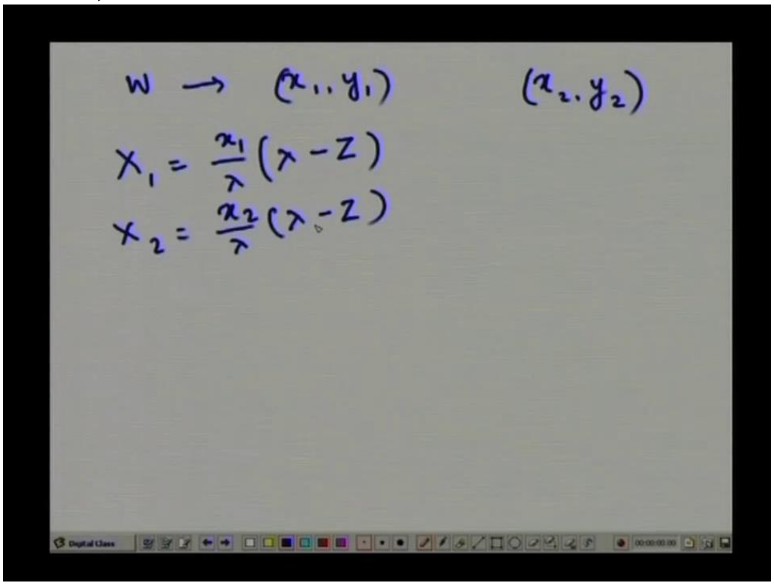
and the right camera Now by applying inverse perspective transformation we find that the equation of straight line with respect to left camera on which point w will lie is given by the equation say x 1 is equal to x 1 by lambda into lambda minus z. Similarly with respect to the right camera the equation of the straight line on which the

(Refer Slide Time 08:36)



same point w will exist is given by x 2 is equal to x 2 by lambda into lambda minus z

(Refer Slide Time 08:50)

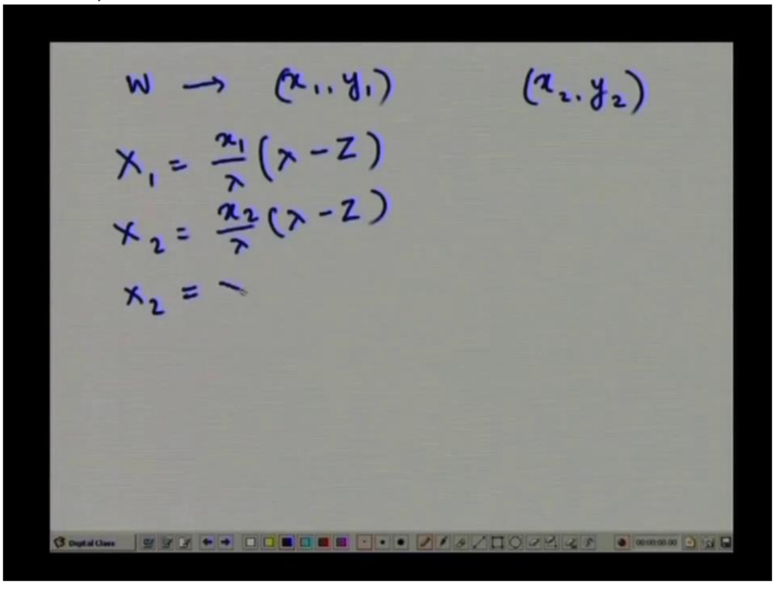


where this capital X 1 is the x coordinate of the point w with respect to the camera coordinate of camera 1, and capital X 2 is the x coordinate of the 3D point w with respect to

(Refer Slide Time 09:07)

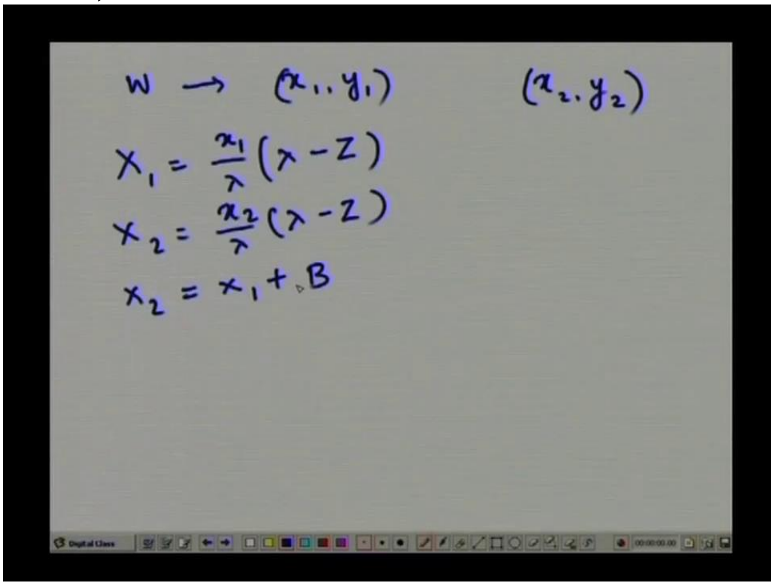
the camera coordinate of the second camera. Now recollect the figure that we have shown that is the arrangement of the camera, where the cameras are displaced by the displacement b. So with respect to that camera arrangement we can easily find out that the value of x 2 will be simply

(Refer Slide Time 09:29)



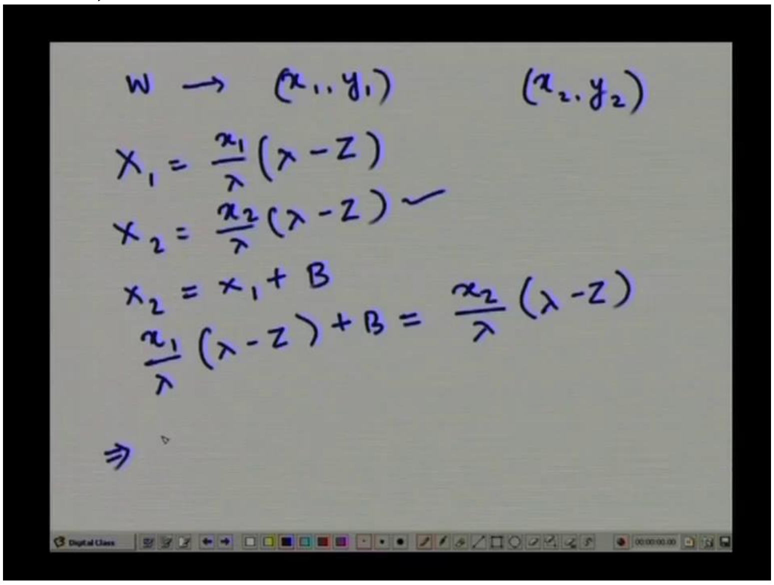
x 1 plus the displacement b Now if I replace

(Refer Slide Time 09:36)



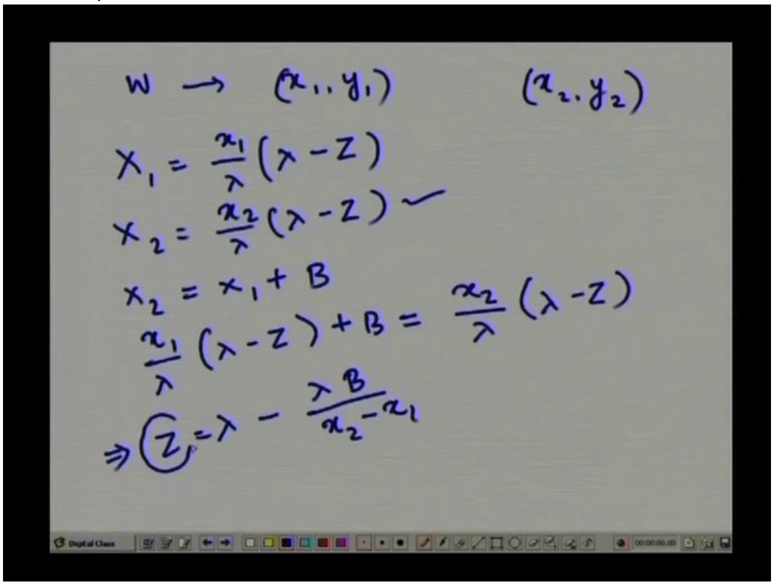
this value of x 2 which is x 1 equal to b in this particular equation then I get a set of equations which gives x 1 by lambda into lambda minus capital Z plus b which is equal to x 2 by lambda into lambda minus z. And from this I get an equation

(Refer Slide Time 10:09)



of the form z equal to lambda minus lambda times b divided by x 2 minus x 1. So find that this z is the z coordinate

(Refer Slide Time 10:25)

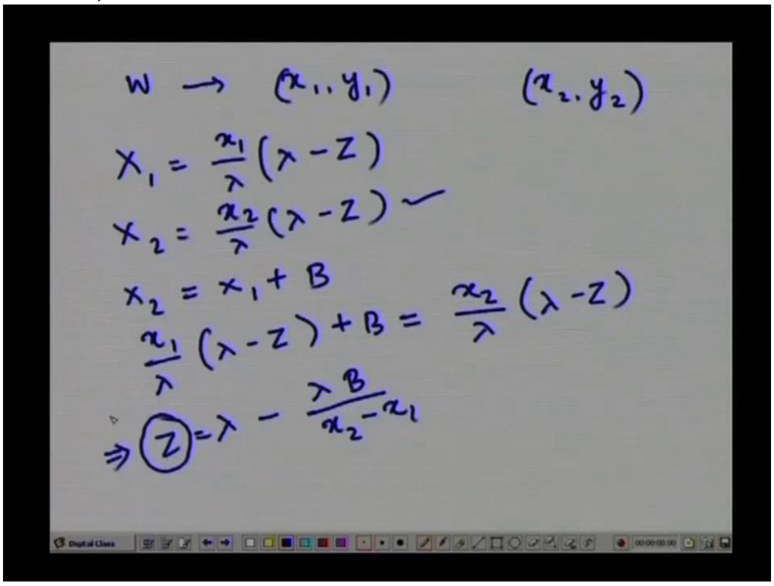


of the 3D point w with respect to the coordinate

(Refer Slide Time 10:31)

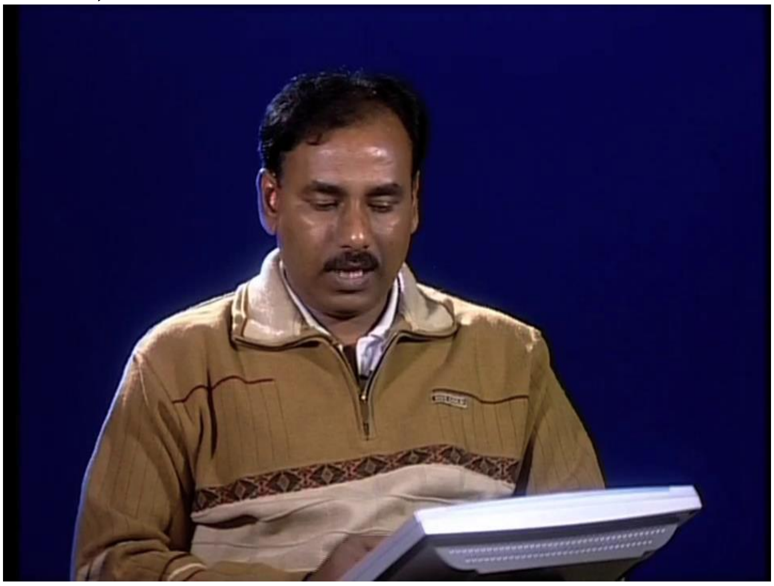
coordinate system of the first camera, it is same as the coordinate system, the z value with respect to coordinate system of the second camera. It is also the z value with respect to the 3D world coordinate system. So that means that it gives me what is

(Refer Slide Time 10:48)



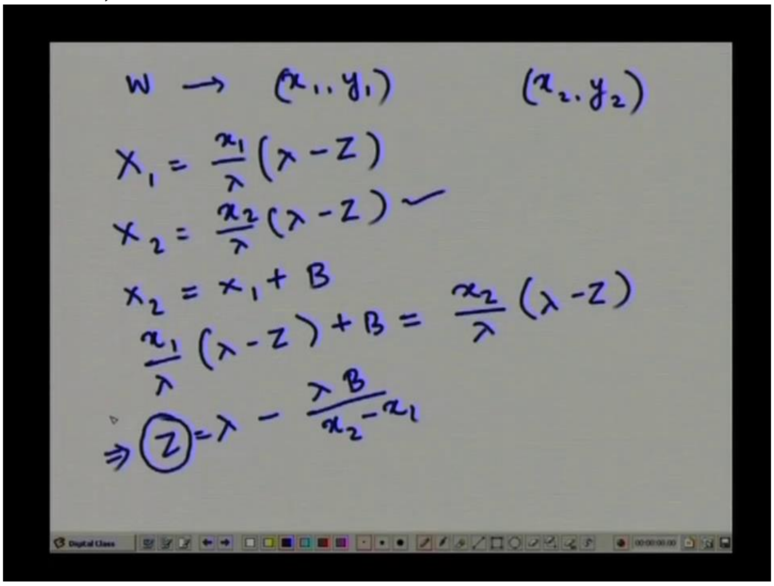
the z value of the 3D point for which the left image point was x 2 y 2

(Refer Slide Time 10:57)



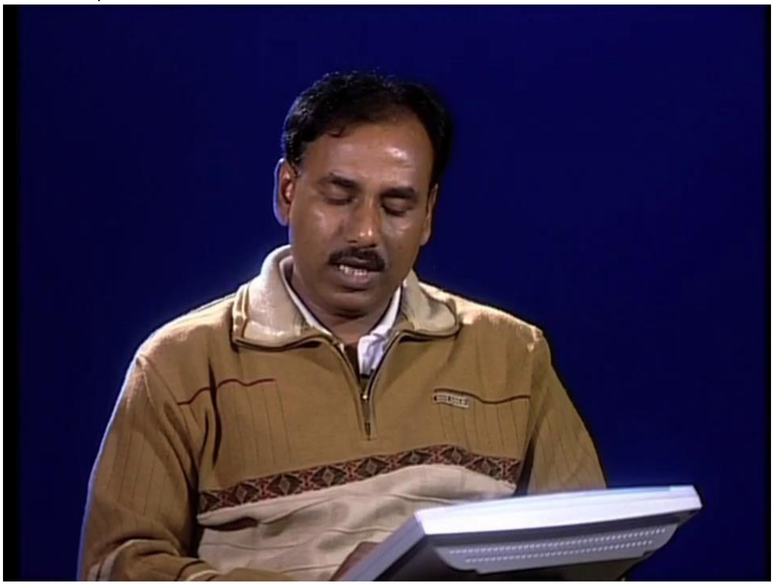
x 1 y 1 and the right image point was x 2 y 2 and I can estimate this value of z from the knowledge

(Refer Slide Time 11:05)



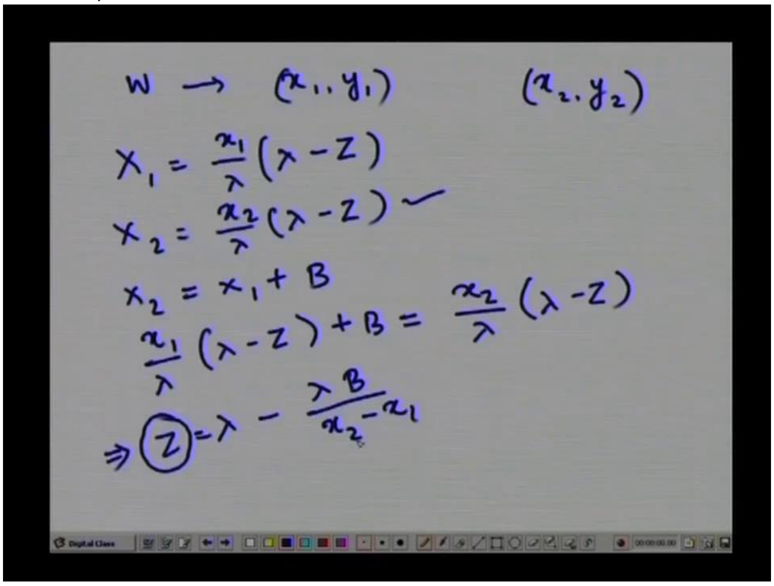
of the wavelength lambda, from the knowledge of displacement between two cameras which is b and the, from the knowledge

(Refer Slide Time 11:14)



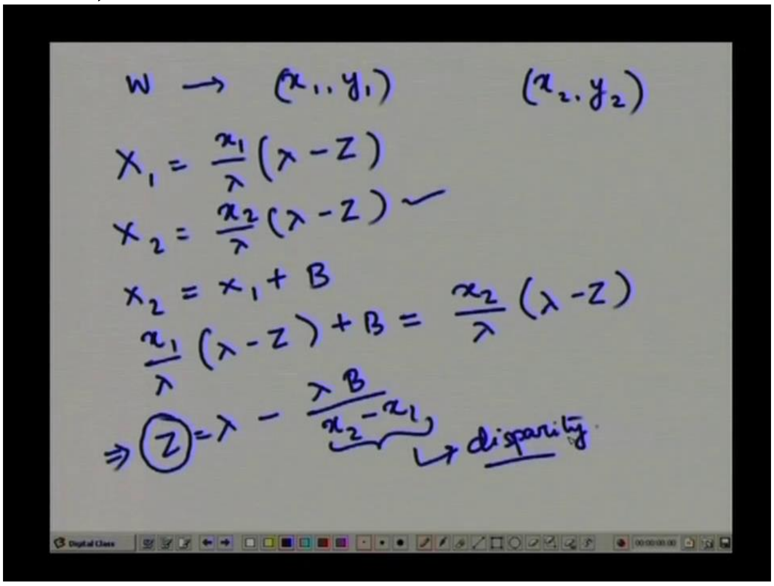
of the difference of the x coordinates that x 2 minus x 1 in the left camera in the left image and the right image. So this

(Refer Slide Time 11:23)



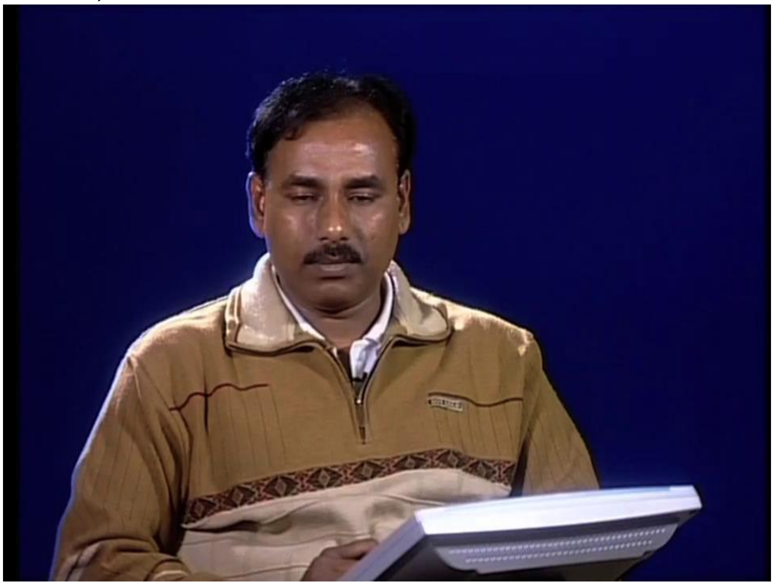
x 2 minus x 1 this term, this particular quantity is also known as disparity. So if know this disparity for a particular point

(Refer Slide Time 11:40)



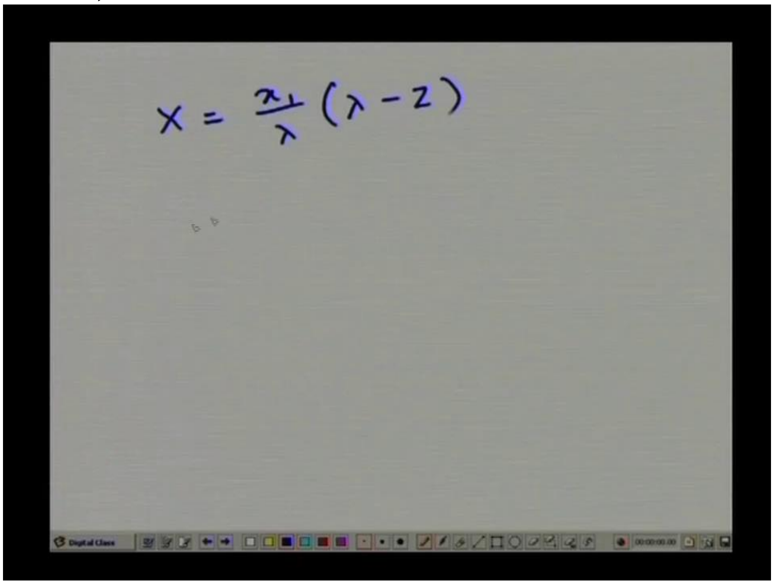
in the left image and the right image, I know the lambda that is focal length of the camera

(Refer Slide Time 11:46)



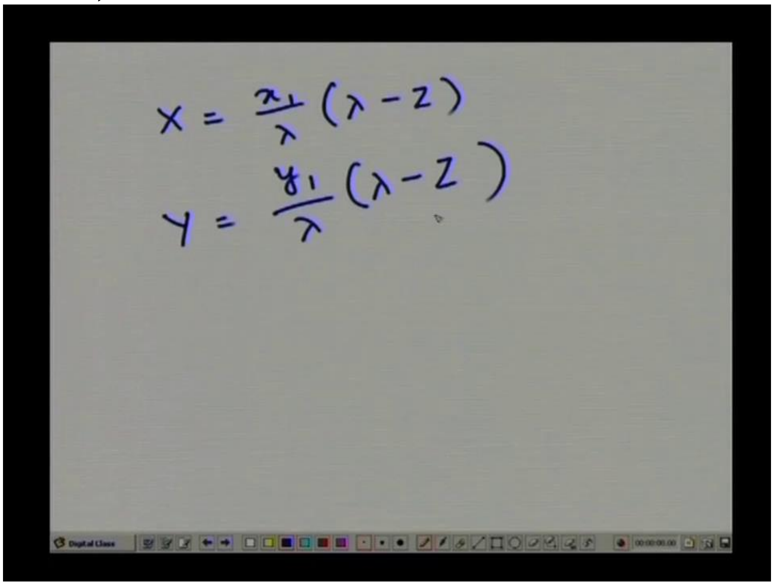
and I know the displacement between the two cameras, I can find out what s the corresponding depth value that is z. And once I know this depth value I can also find the x coordinate and y coordinate of the 3D point w with respect to the 3D world coordinate system for which we have already seen the equations are given by x equal to x 1 by lambda into lambda minus z

(Refer Slide Time 12:18)



and y equal to y 1 by lambda into lambda minus z. So first we have computed

(Refer Slide Time 12:26)

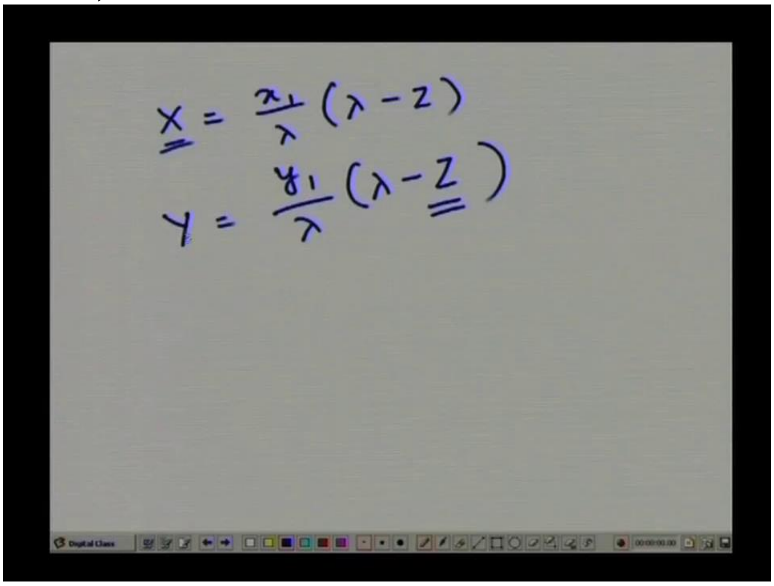


the value of z from the knowledge of disparity, camera focal length

(Refer Slide Time 12:32)

and the displacement between the cameras and then from this value of z and the image coordinates in say left image that is x 1 y 1 I can find out what is the x value x coordinate value

(Refer Slide Time 12:47)



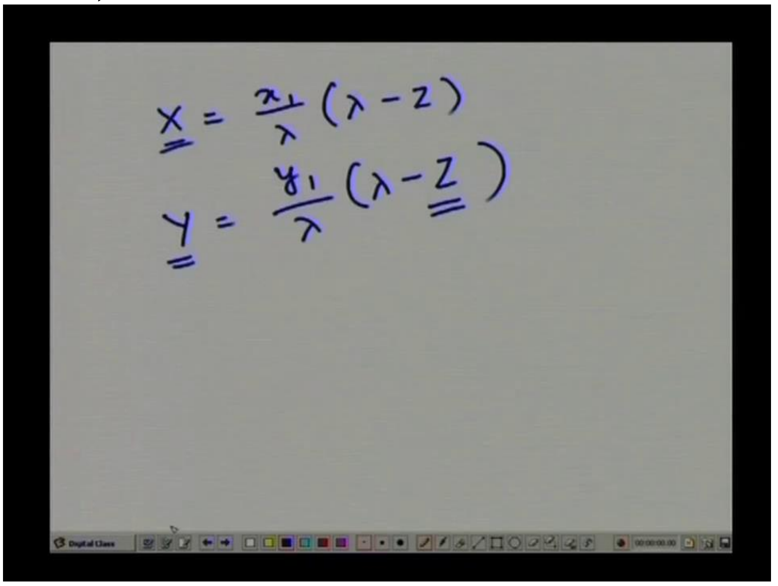
and y coordinate value of that particular 3D point.

Now in this you find that very, very important computation is

(Refer Slide Time 12:57)

that given a point in the left image, what will be corresponding point

(Refer Slide Time 13:03)

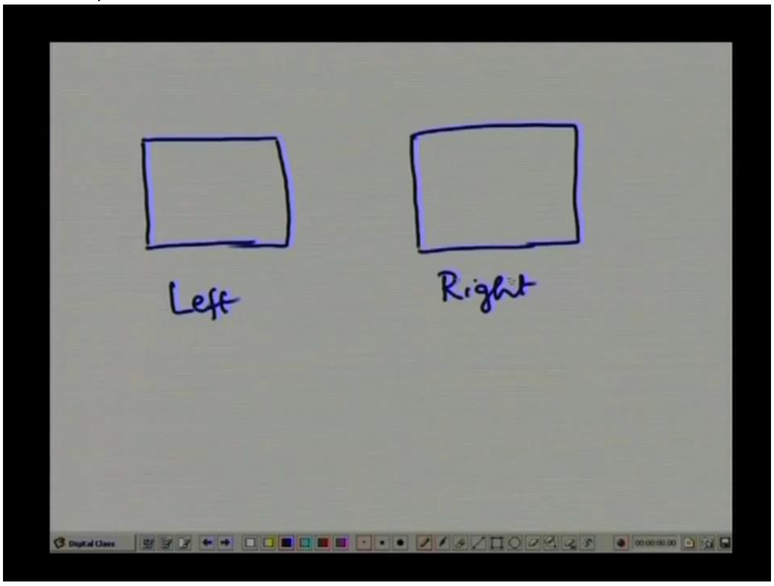


in the right image So this is the problem which is called as stereo correspondence problem. So in

(Refer Slide Time 13:13)

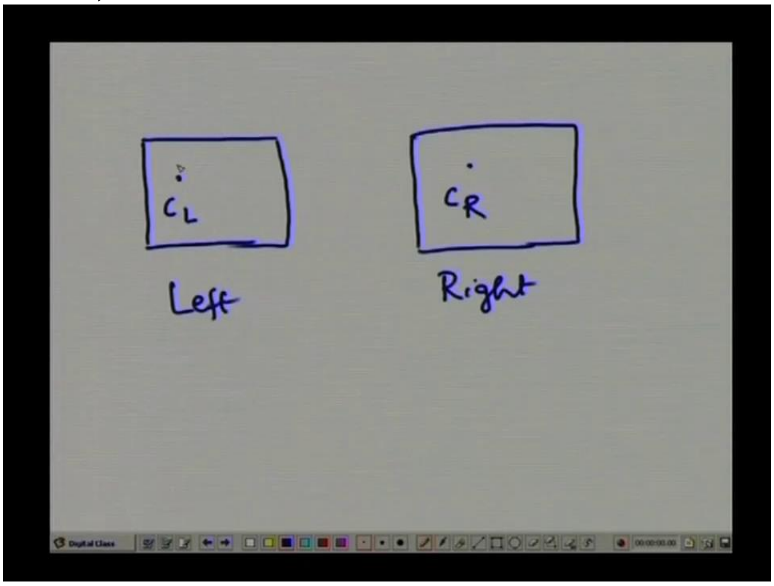
today's lecture we are not going to deal with the details of the stereo correspondence problem that is how do you find out a point in the left image and the corresponding point in the right image. But today what we will discuss is about the complexity of this correspondence operation. So our problem is like this. We have a left image and we have a right image. So this is the left image and this is the right image. So

(Refer Slide Time 13:50)



if I have a point say c l in the left image I have to find a point c r in the right image which corresponds to c l. And once I do this here

(Refer Slide Time 14:05)



I find out what is the coordinate, image coordinate of this point c l which is say 1 by 1 and which is image coordinate of this point c r which is 2 by 2 so once I know these

(Refer Slide Time 14:20)

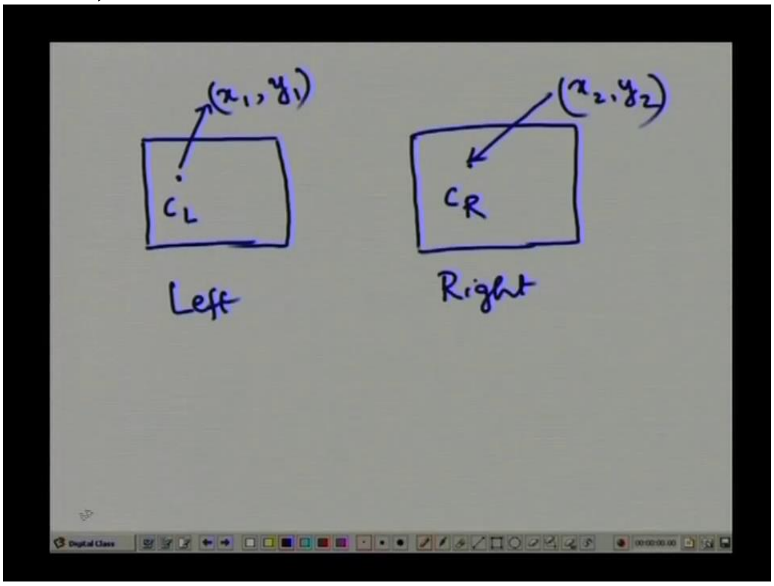
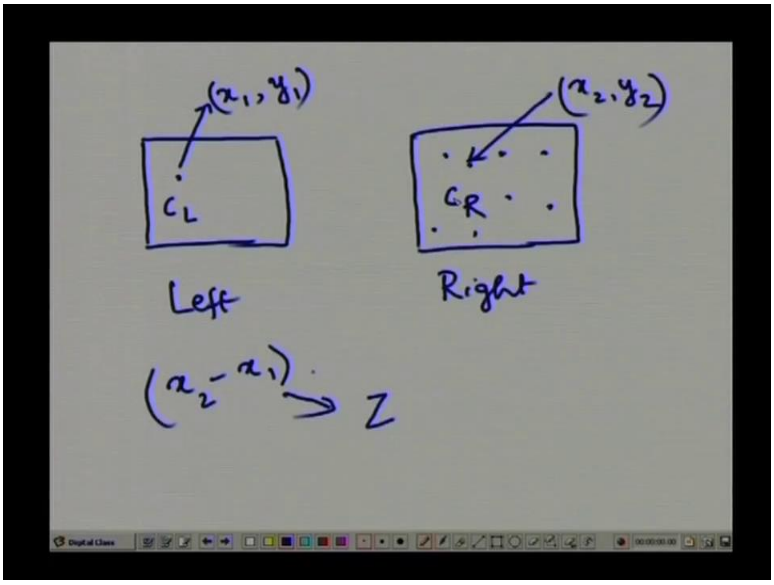


image coordinates I can compute x 2 minus x 1 which is the disparity and then this x 2 minus x 1 is used for the computation of the z value z. Now what about the complexity of this search operation? Say I identify a particular point say c l in the left image, then a corresponding

(Refer Slide Time 14:41)

point c r in the right image may appear anywhere in the right image. So if I have

(Refer Slide Time 14:50)

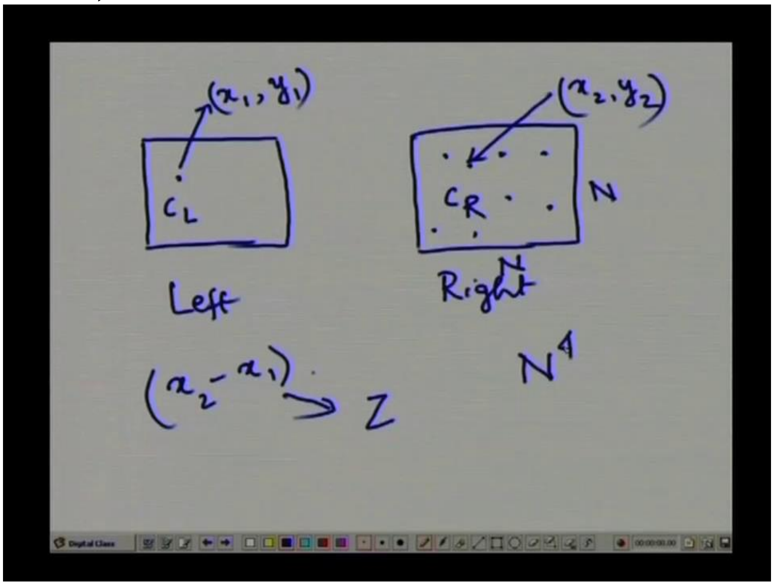


the images whose dimensions are of order n by n that means I have n number of rows and n number of columns then you find that for every point in the left image I have to search n square

(Refer Slide Time 15:04)

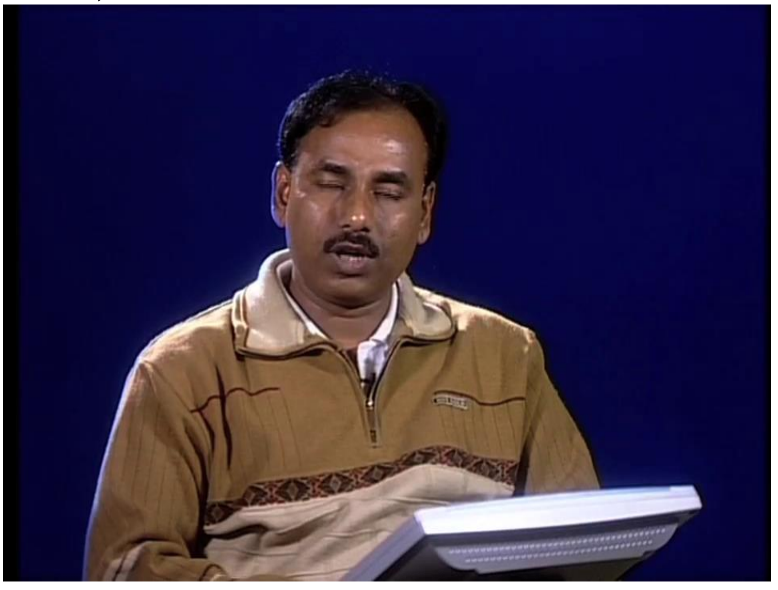
number of points in the right image And because there are n square number of points in the left image, in the worst case I have to search for n to the power 4 number of points

(Refer Slide Time 15:16)



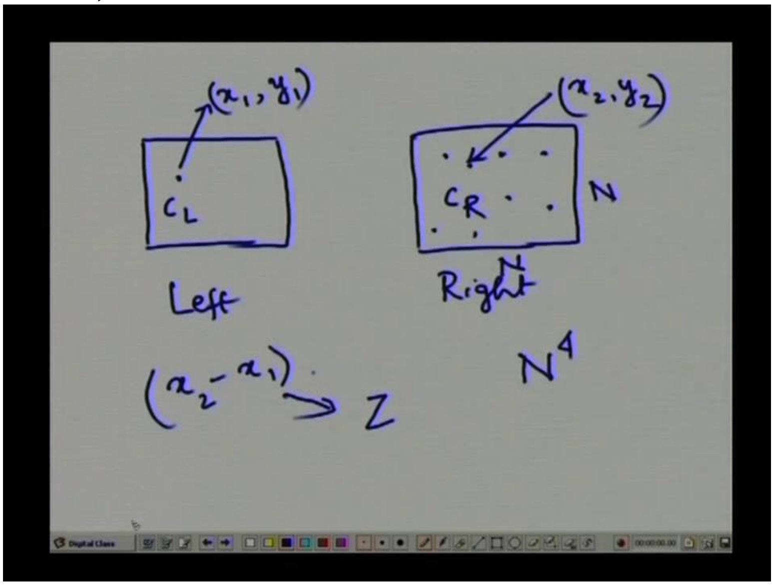
to find out the correspondence

(Refer Slide Time 15:18)



correspondence for every point in the left image and the corresponding point in the right image so this is a massive computation; so, how to reduce this computation? Fortunately

(Refer Slide Time 15:32)

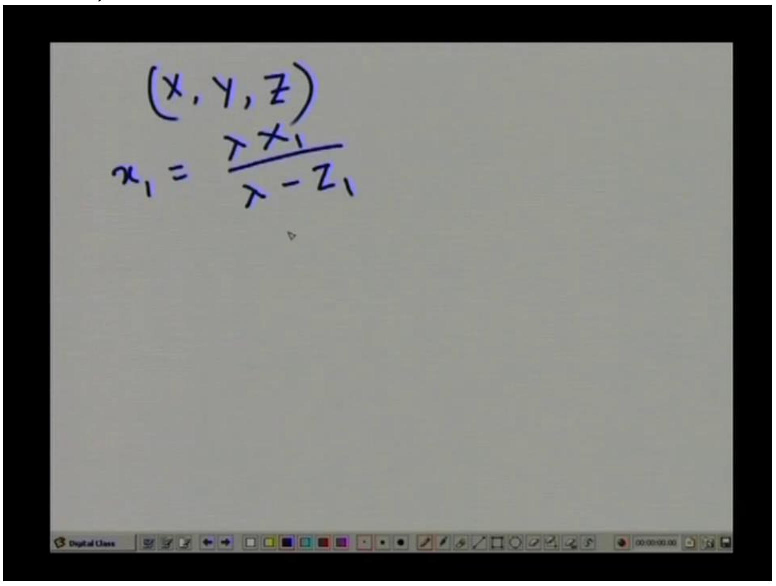


the imaging geometry that we have used, that helps us in reducing the computation that we will be doing. So find that

(Refer Slide Time 15:46)

for the point x y z in the 3D space the corresponding left image is given by x 1 is equal to lambda times capital X 1 divided by lambda minus capital Z 1. So I assume

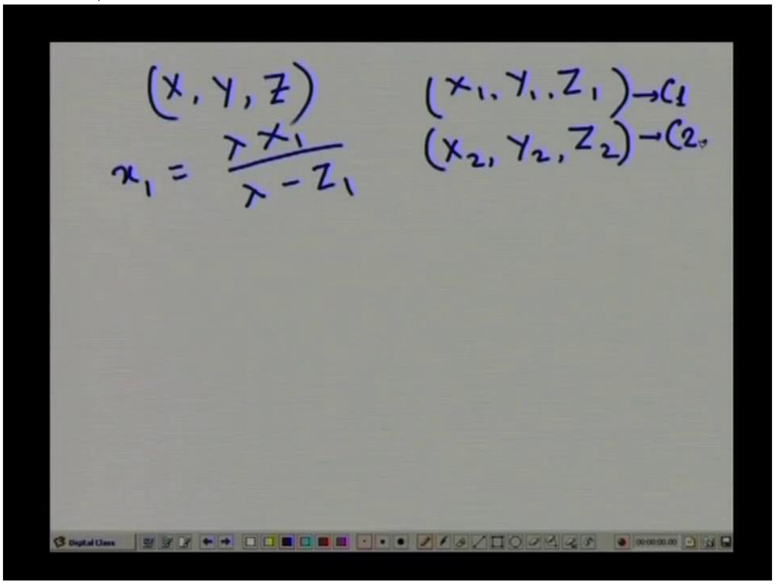
(Refer Slide Time 16:09)



that capital X 1 capital Y 1 and capital Z 1, they are the image coordinates of the first camera. And I also assume that capital X 2 capital Y 2 and capital Z 2; they are the image coordinates for the second camera. So this is for camera 1 and this is for camera 2.

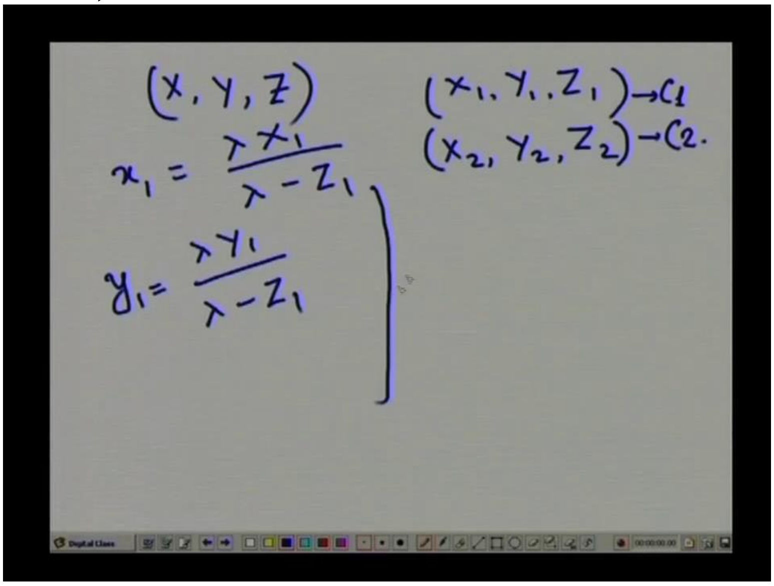
So with

(Refer Slide Time 16:35)



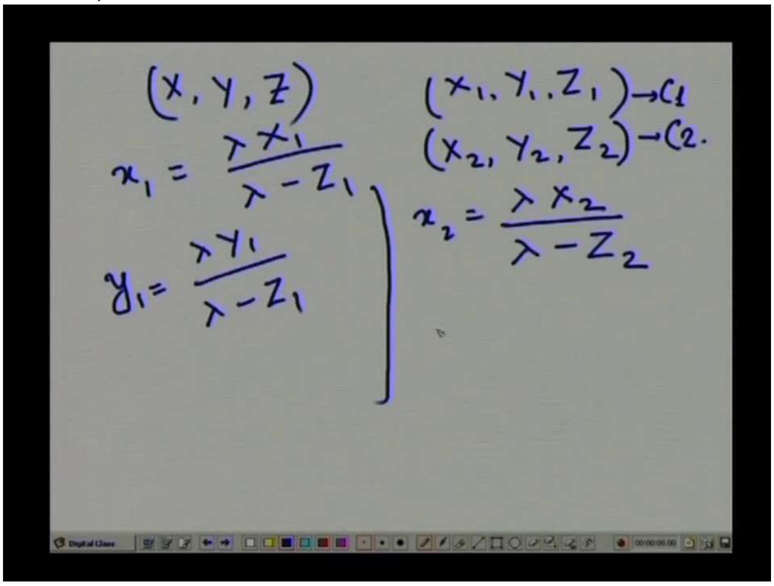
respect to camera 1, the value of x 1, image point x 1 is given by lambda times capital X 1 divided by lambda minus capital Z 1. Similarly y 1, the y coordinate in the first image is given by lambda times capital Y 1 divided by lambda minus z 1. Now with respect to the second image

(Refer Slide Time 17:02)



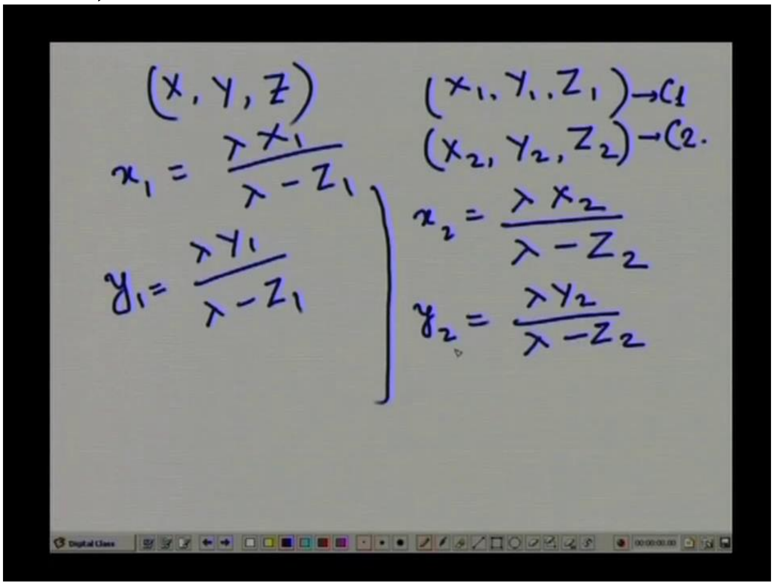
the image coordinate x 2 is given by lambda times capital X 2 divided by lambda minus capital Z 2

(Refer Slide Time 17:14)



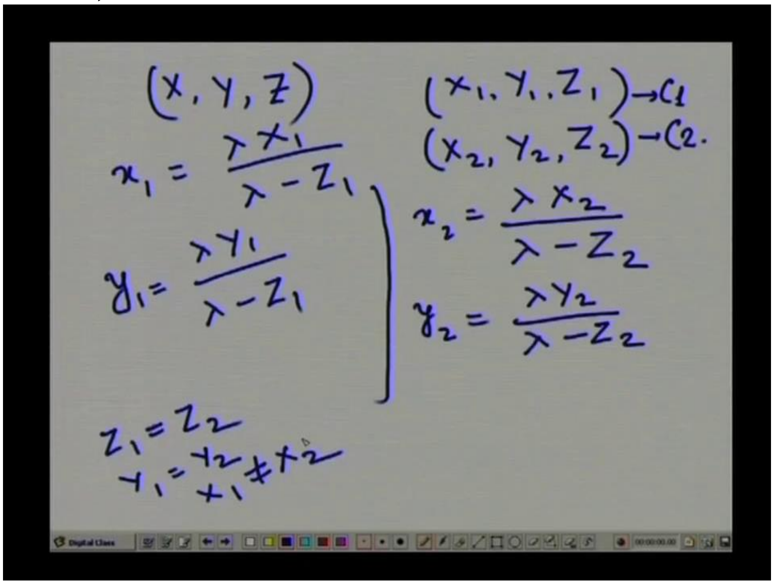
Similarly y 2 is also given by lambda times y 2 divided by lambda minus capital Z 2.

(Refer Slide Time 17:25)

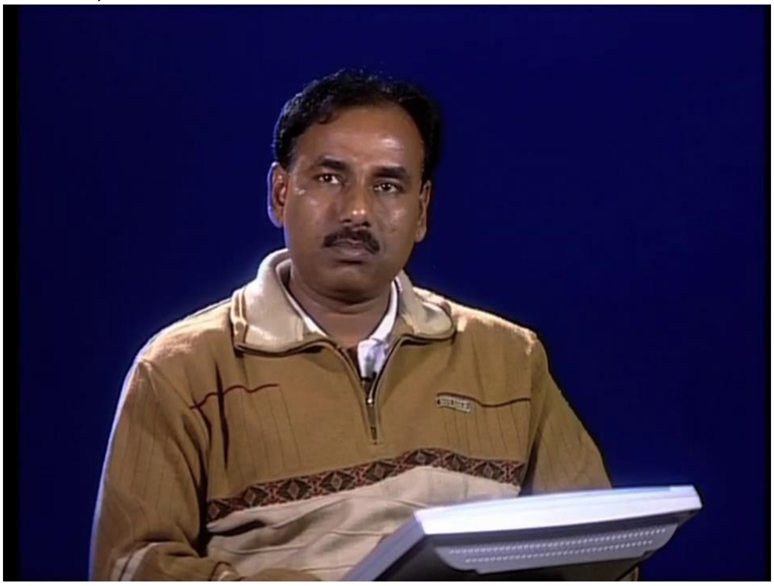


Now we find that the imaging system or the imaging setup that we have used, in that we have said, we have seen that capital Z 1 is equal to capital Z 2, capital Y 1 is equal to capital Y 2 but capital X 1 is not equal to capital X 2. This is because

(Refer Slide Time 17:52)

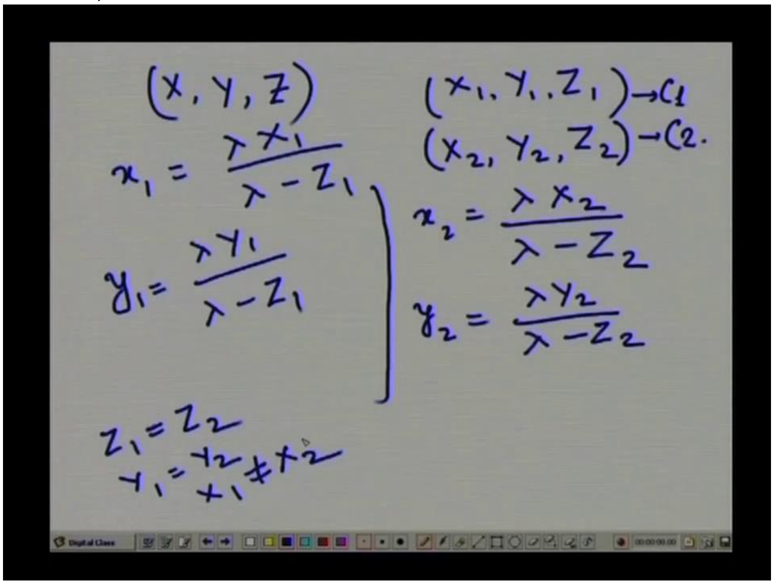


(Refer Slide Time 17:53)



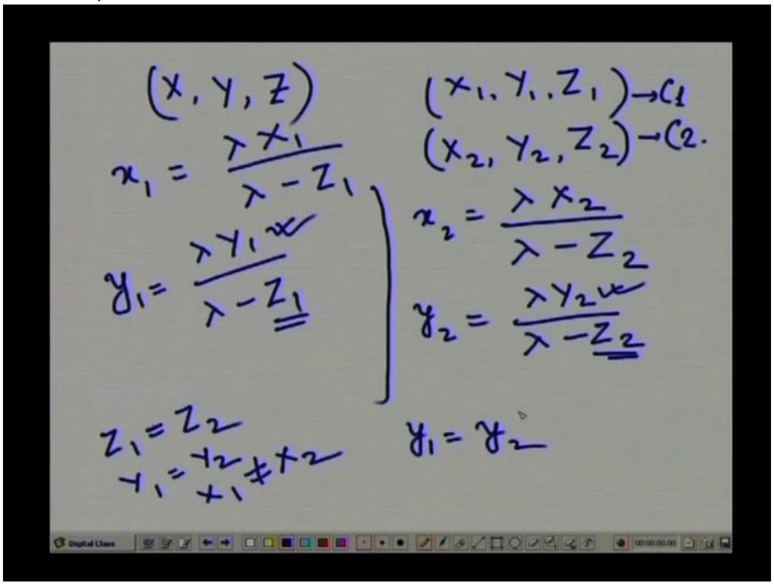
the two cameras are displaced only in the x direction. They don't have any displacement in the y direction. Neither they have any displacement in the z direction. So for both the camera coordinate systems the x coordinate, sorry the z coordinate and the y coordinate value

(Refer Slide Time 18:10)



for both the cameras will be same whereas the x coordinate will be different. So by applying that, since z 1 is equal to z 2 and y 1 and is also equal to y 2, so you find that among the image coordinates on the two images image 1 and image 2, y 1 will be equal to y 2. So what does

(Refer Slide Time 18:40)

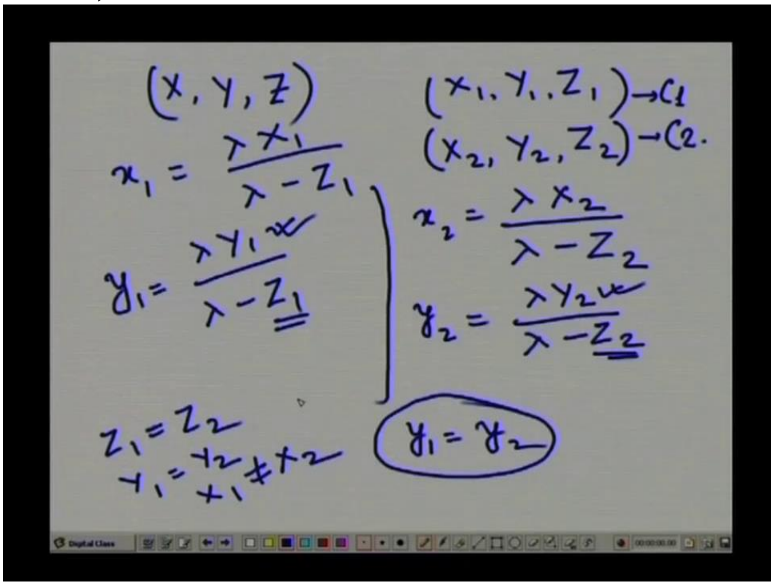


this mean? This means that whatever is the x 1 y 1 value

(Refer Slide Time 18:46)

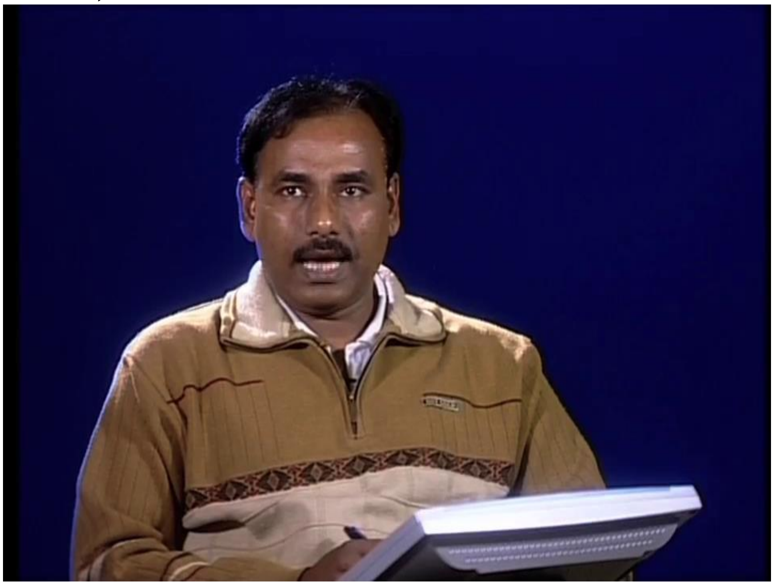
of point c l in the left image, the corresponding right image point c r will have a different x value, x coordinate value but it will have the same y coordinate value. That means two corresponding points, image points must lie on the same row. So if I pick up c l belonging to row I in the left image, the corresponding point c r in the right image will also belong to the same row I. So by this for a given point I don't have to search the entire right image to find out the correspondence but I will simply search that particular row to which c l belongs, that particular row in right image to find out a correspondence. So this saves a lot of time for searching for correspondence between a point in the left image

(Refer Slide Time 19:41)



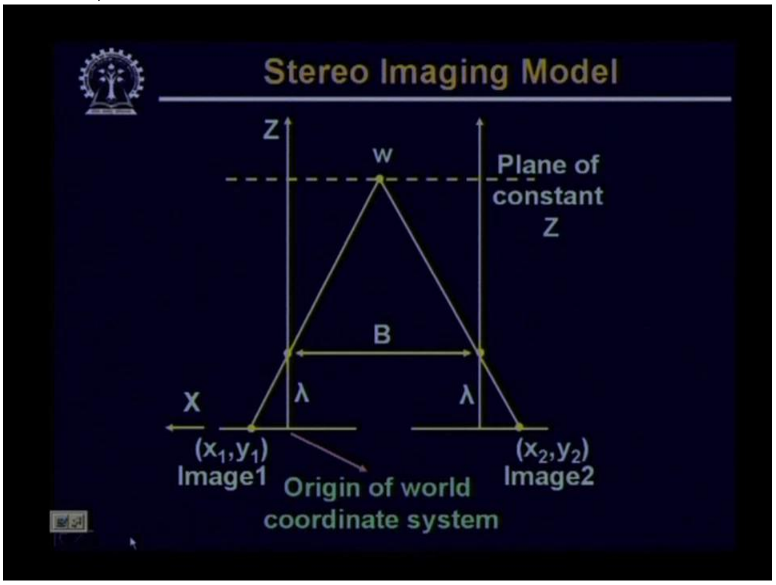
and the corresponding point in the right image. So till now we have discussed that how using two different cameras

(Refer Slide Time 19:54)



and having a stereo imaging setup we can find out the 3D world coordinates of the points which have a point, an image point in the left image and the corresponding point in the right image. But by studying this stereo imaging setup you can find out that it is, it may not always be possible to find out a point in the right image for every possible point in the left image. So there will be a certain region, there will be a certain region in three-dimensional space where for which space for all the points in that space I will have image points both in the left image and the right image but for any point outside that region I will have points only in one of the images, either in the left image or in the right image but I cannot have points in both the images. And unless I have points in both the images I cannot estimate the three-dimensional x y z coordinate of those points. So till now we have seen that using a single camera

(Refer Slide Time 21:09)



camera I cannot estimate the depth value of a point in 3D but if I have

(Refer Slide Time 21:16)

2 cameras and using stereo setup I can estimate the depth value of the three-dimensional points. Thank you.