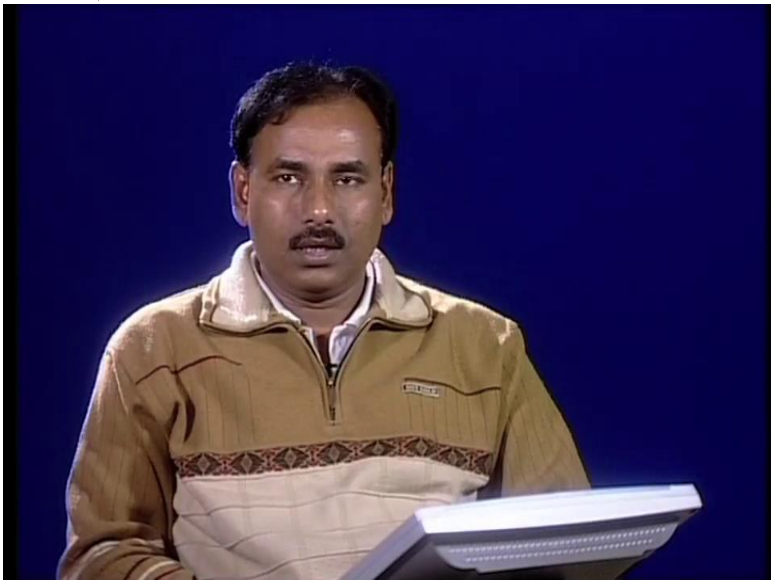
Digital Image Processing

Prof. P. K. Biswas
Department of Electronics and Electrical Communications Engineering
Indian Institute of Technology, Kharagpur

Module 03 Lecture Number 13

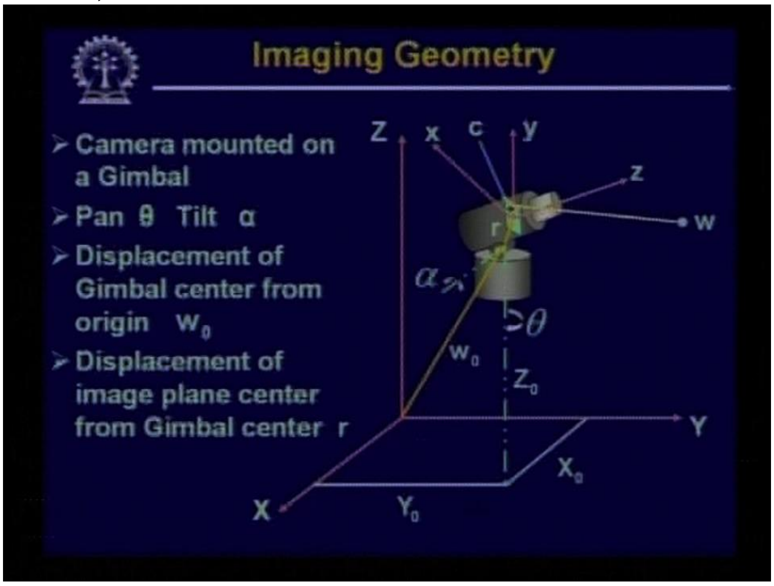
Image Geometry - 1

(Refer Slide Time 00:19)



Hello, welcome to the video lecture series on Digital Image Processing. Indentify the point w which maps to point x naught y naught in the image plane. Now till now, all the discussions that we had done, for all these discussions that we had assumed that the image coordinate system and the camera coordinate system, they are perfectly aligned. Now let us discuss about a general situation where the image coordinate system and the camera coordinate system, they are not perfectly aligned. So here we assume that the

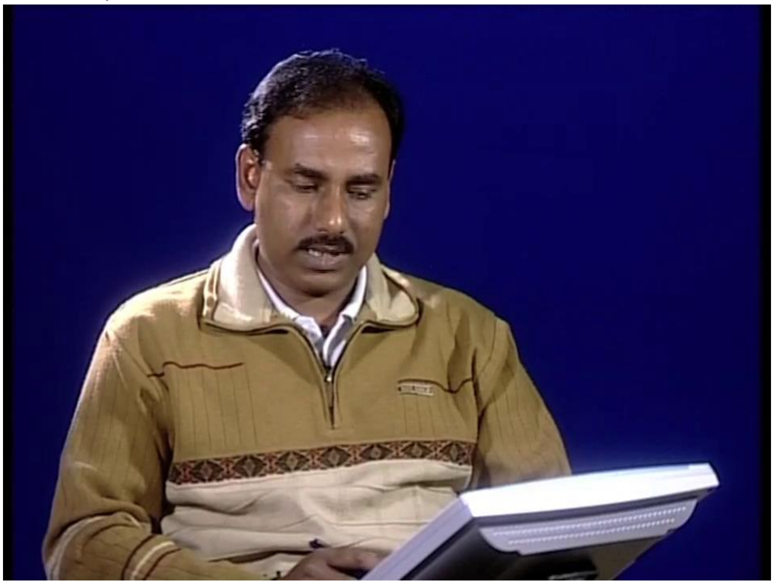
(Refer Slide Time 01:00)



camera is mounted on a gimbal. So if you mount the camera on the gimbal then using the gimbal the camera can be given a pan of angle theta, it can also be given a tilt by an angle alpha. So you remember that pan is the rotation around z axis and the tilt is the rotation around x axis. We also assume that the gimbal center is displaced from the 3D world coordinate origin 0 0 0 by a vector w naught which is equal to x naught y naught z naught and finally we also assume that the camera center or the center of the imaging plane is displaced from the gimbal center by a vector r which will have components say r 1, r 2 and r 3 in the x, y and z direction of the 3D world coordinate system.

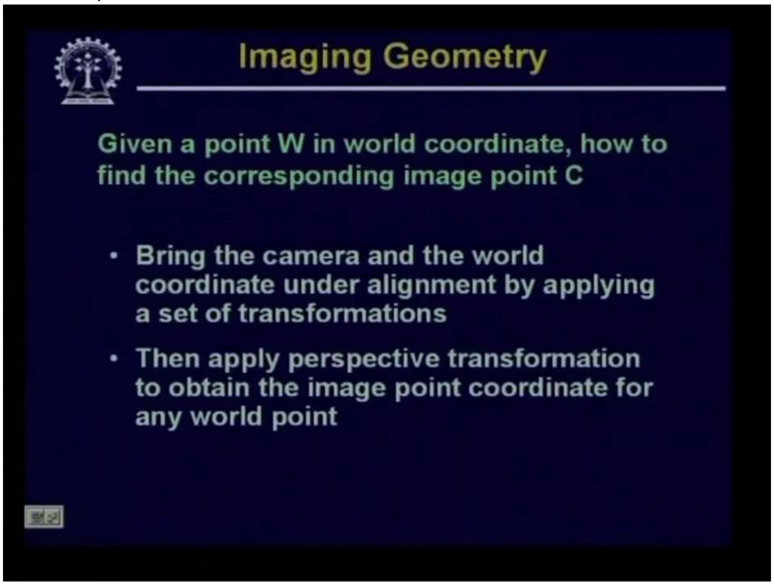
Now here our interest is, given such a type of imaging arrangement, now if we have a 3D world, a point in the 3D world coordinate w what would be the camera coordinate, what would be the image point c to which this world point w will be mapped? So this is a general situation. And let us see how we can obtain the solution to this particular problem that for this generalized imaging setup, for a world point w what will be the corresponding image point c?

(Refer Slide Time 02:40)



So

(Refer Slide Time 02:43)

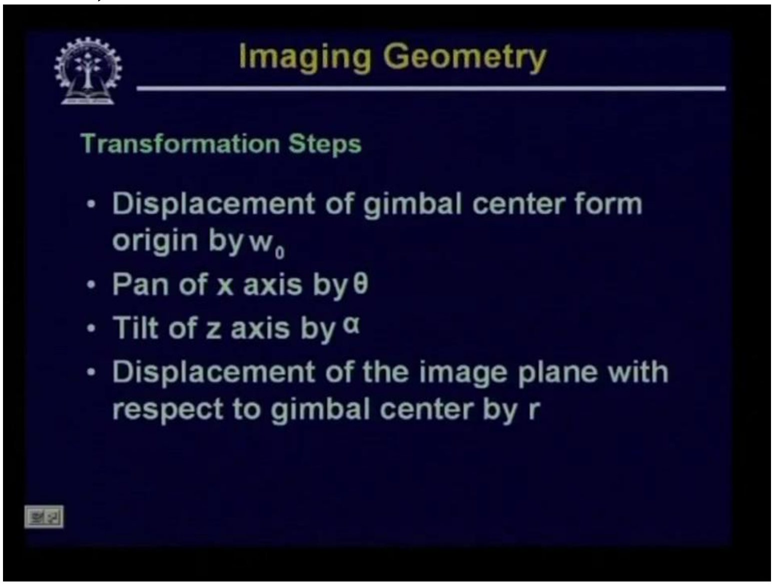


So the steps would be like this. Since our earlier formulations were very simple in which case we had assumed that both the camera coordinate system and the 3D world coordinate system, they are perfectly

(Refer Slide Time 02:56)

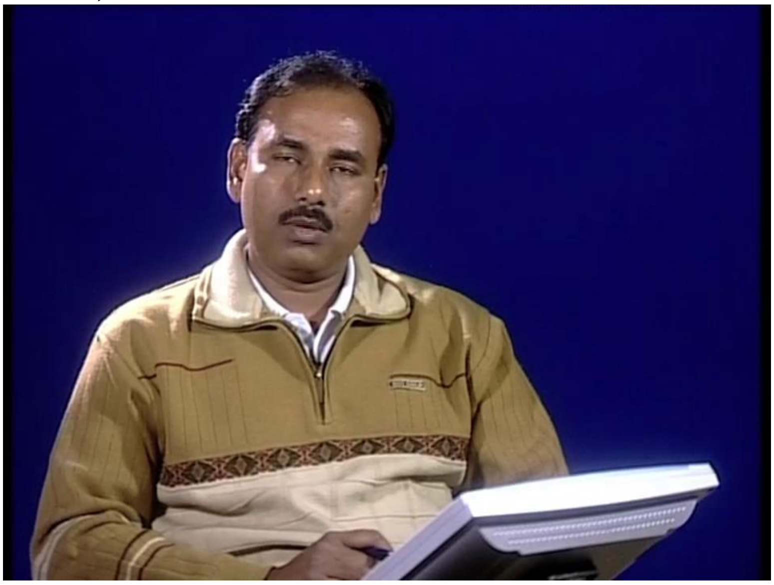
aligned, in this generalized situation we will also try to find out a set of transformations if applied one after another will bring the camera coordinate system and the world coordinate system in perfect alignment. So once that alignment is made, then we can apply the perspective transformation to the transformed 3D world points and this perspective transformation to the transformed 3D world points give us the corresponding image coordinate of the transformed point w. So what are the transformation steps that we need in this particular case?

(Refer Slide Time 03:40)



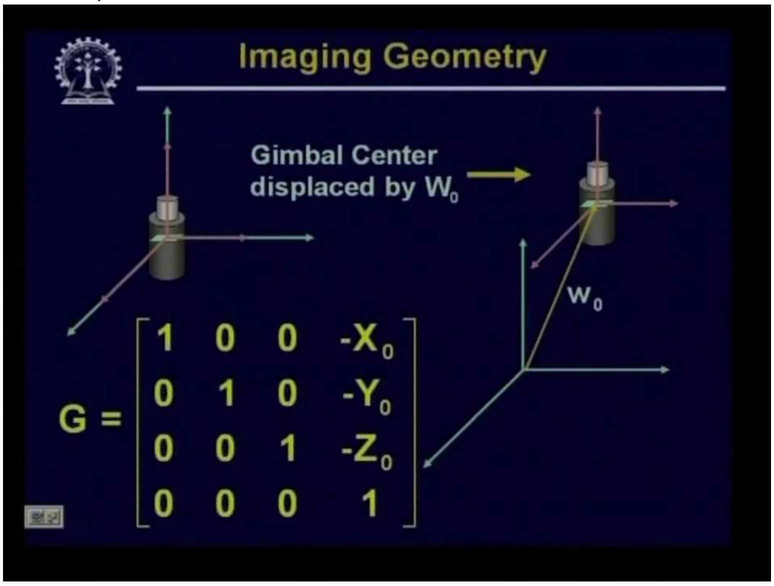
So the first step is we assume that the image coordinate system and the 3D world coordinate system, they are perfectly aligned. So from this we displace the gimbal center from the origin by the vector w naught and after displacing the gimbal center from the origin by w naught, we pan along x axis by an angle theta followed by tilt around z, tilt of z axis by angle alpha which will be followed by the final displacement of the image plane with respect to gimbal center by the vector r. So we have 4 different transformation steps

(Refer Slide Time 04:29)



which are to be applied one after another and these transformation steps will give you the transformed coordinates of the 3D world point w. So let us see how this transformation is to be applied one after another.

(Refer Slide Time 04:43)



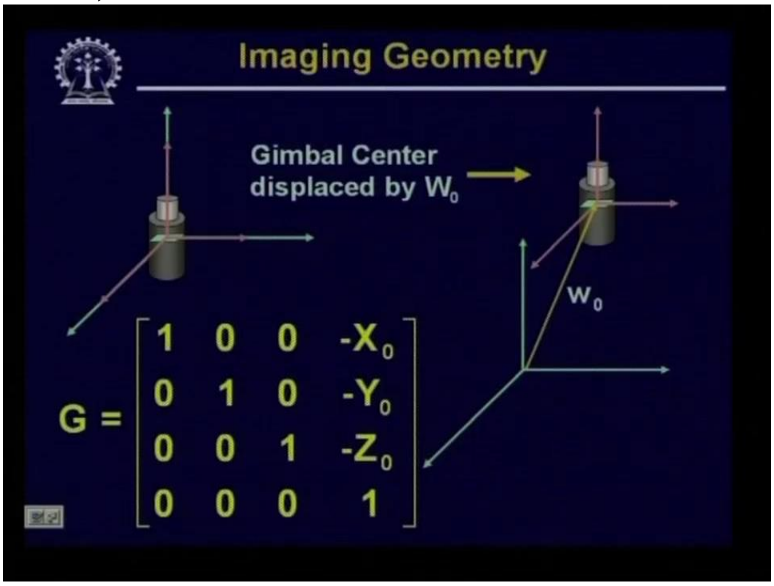
So here on the left hand side we have shown a figure where the camera coordinate system and the world coordinate system are perfectly aligned. Now from this alignment if we give a displacement by a vector w naught to the gimbal center then the camera will be displaced as shown on the right hand side of the figure where you find that the center is displaced by vector w naught. You remember that if I displace the camera center

(Refer Slide Time 05:19)

by vector w naught then all the world coordinates, all the world points will rep displaced by a vector minus w naught with respect to the camera. Now you just recollect that when we tried to find out the image point of a 3D world point then the image point, the location of the image point is decided by the location of the 3D world point with respect to the camera coordinate. It is not with respect to the 3D world coordinate.

So in this case, also after the set of transformations we have to find out what are the coordinates of the 3D world point with respect to the camera coordinate system where originally the coordinates of the 3D world point are specified with respect to the 3D world coordinate systems. So here as we displace the camera center by vector w naught, so all the world coordinate points, all the world points will be displaced by the vector which is negative of w naught that is by minus w naught and if w naught has components x naught along x direction, y naught along y direction and z naught along z direction so the corresponding transformation to the 3D points will be minus x naught, minus y naught and minus z naught. And we have seen earlier that if a 3D point is to be displaced by minus x naught, minus y naught, minus z naught then in the uniform representation the corresponding transformation matrix for this translation is given by g equal to

(Refer Slide Time 07:08)



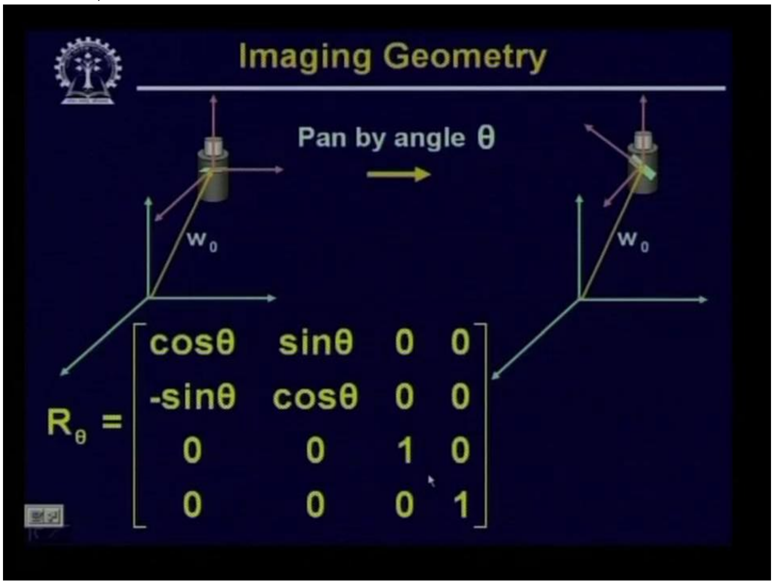
1 0 0 minus x naught, 0 1 0 minus y naught, 0 0 1 minus z naught and then 0 0 0 1 So this is a transformation matrix which translates all the world coordinates, all the world points by vector minus x naught minus y naught minus z naught and this transformation is now

(Refer Slide Time 07:33)

with respect to the camera coordinate system.

The next operation as we said

(Refer Slide Time 07:39)

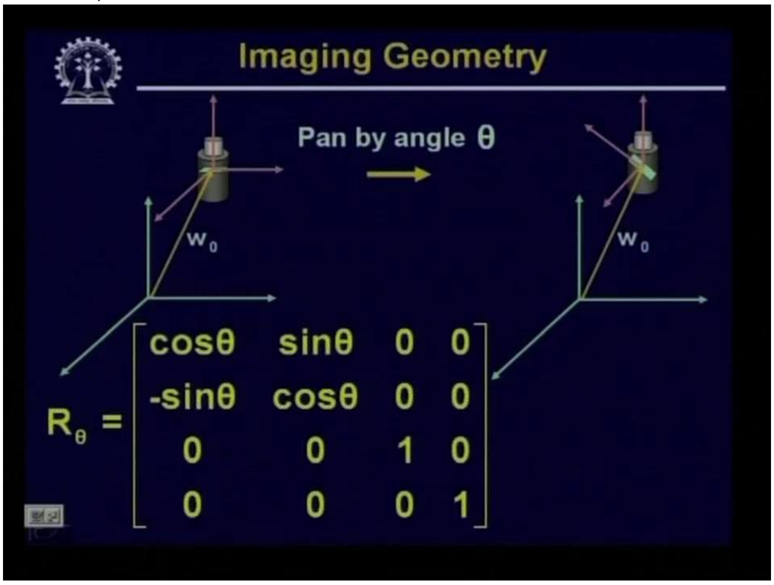


after this displacement, we pan the camera by angle theta. And this panning is done along the z axis. So when we pan along the z axis, the coordinates which are going to change

(Refer Slide Time 07:56)

is the x coordinate and the y coordinate The z coordinate value is not going to change at all And for this panning by an angle theta, again we have seen earlier that the corresponding transformation matrix for rotation theta is given by r theta equal to cosine theta sine theta 0 0

(Refer Slide Time 08:17)



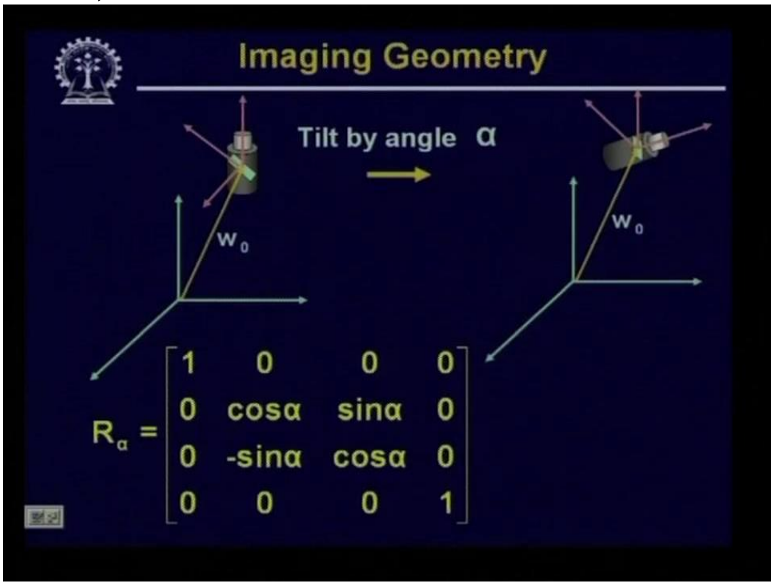
minus sine theta cosine theta 0 0, then 0 0 1 0 and then 0 0 0 1.So when we rotate the camera by an angle theta all the world coordinate points, all the world points will be rotated by the same angle theta by the same angle theta but in opposite direction and that corresponding matrix will be given by this matrix r theta.

So we have completed two steps, first displacement of the camera center with respect to the origin of the world coordinate

(Refer Slide Time 08:54)

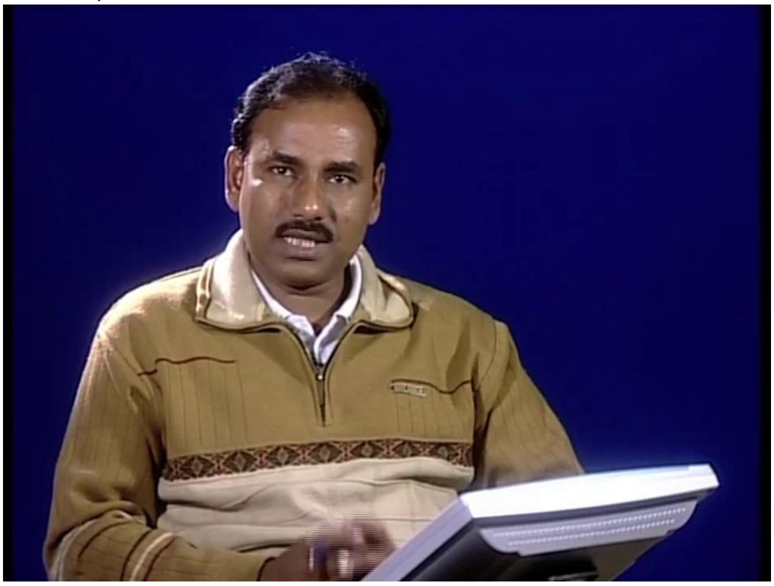
system, then panning the camera by angle theta The third step is, now we have to tilt the camera by an angle alpha. And again we have to find out what is the corresponding transformation for this tilt operation which has to be applied to all the 3D points. So for this tilt

(Refer Slide Time 09:15)



operation by an angle alpha, the corresponding transformation matrix r alpha will be given by 1 0 0 0, 0 cosine alpha sine alpha 0, then 0 minus sine alpha cosine alpha 0 and 0 0 0 1. So you just recollect

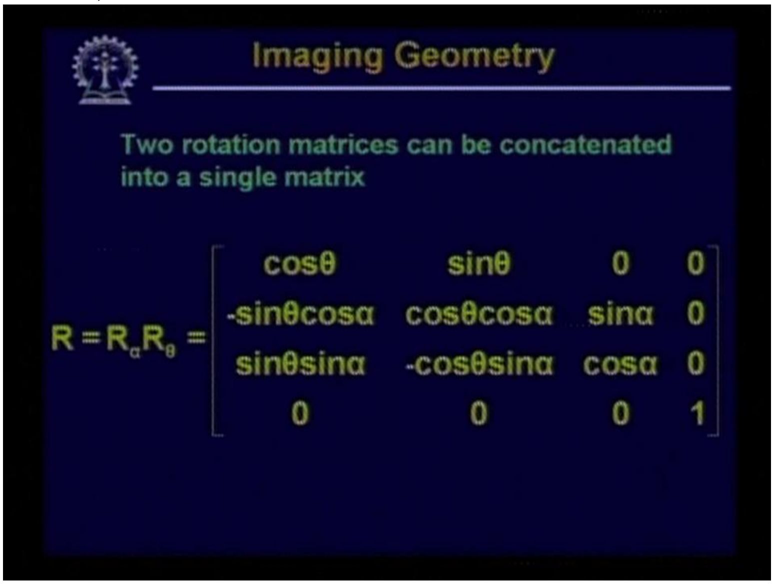
(Refer Slide Time 09:36)



that these are the basic transformations which we had already discussed in the previous class and how these transformations are being used to understand the imaging process

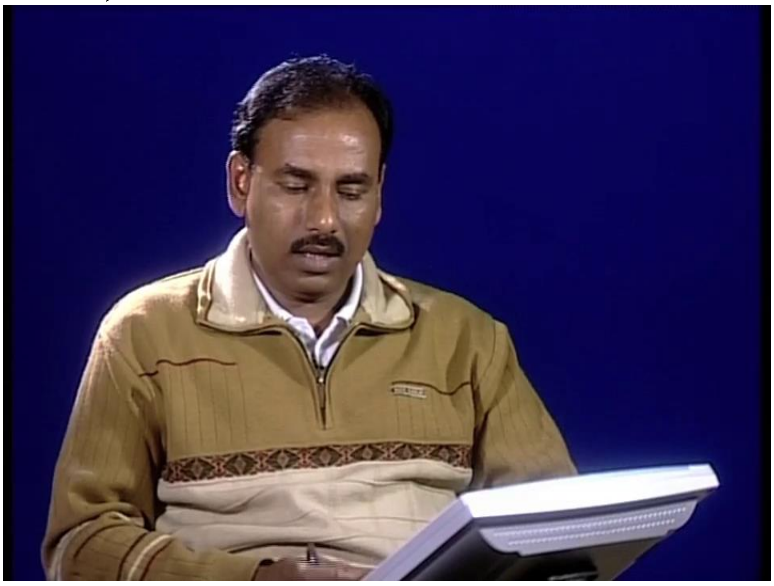
So so far we have applied one displacement and two rotation transformations along r theta and r alpha. Now find that this r theta and r alpha

(Refer Slide Time 09:59)



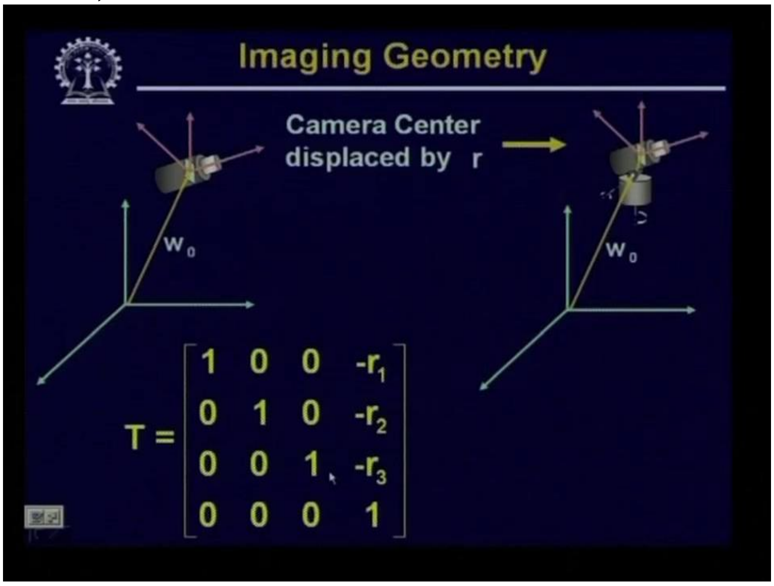
they can be combined into a single rotation matrix r which is equal to r alpha concatenated with r theta and the corresponding transformation matrix r will be given by cosine theta sine theta 0 0, then minus sine theta cosine alpha cosine theta sine alpha, sine alpha 0 then sine theta sine alpha minus cosine theta sine alpha then cosine alpha 0, and then 0 0 0 1.

(Refer Slide Time 10:33)



Then final transformation

(Refer Slide Time 10:37)

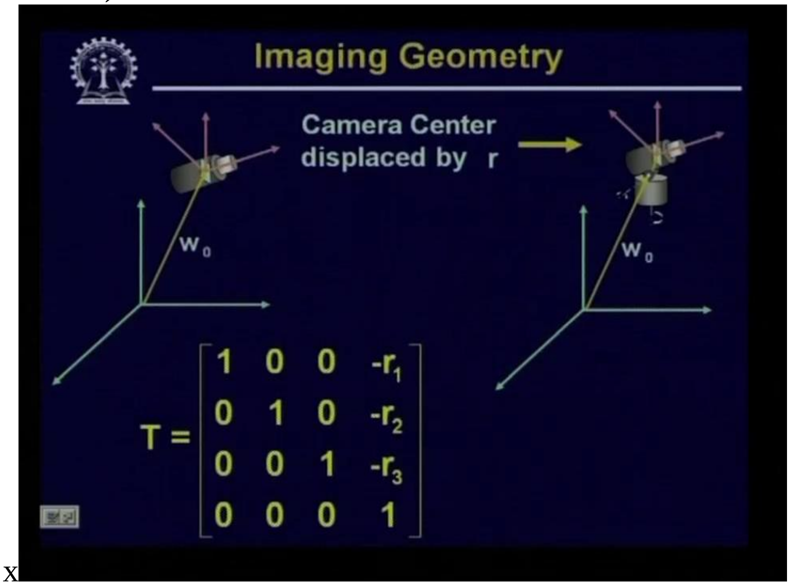


that we have to give to the camera center or the center of the imaging plane from the gimbal center by a vector r and this vector r has the components r 1, r 2 and r 3 along x, y and z directions. And by this transformation, now all the world points

(Refer Slide Time 11:01)

are to be transformed, are to be translated by a vector minus r 1, minus r 2, minus r 3 and the corresponding translation matrix now will be t equal to 1 0 0 minus 1,

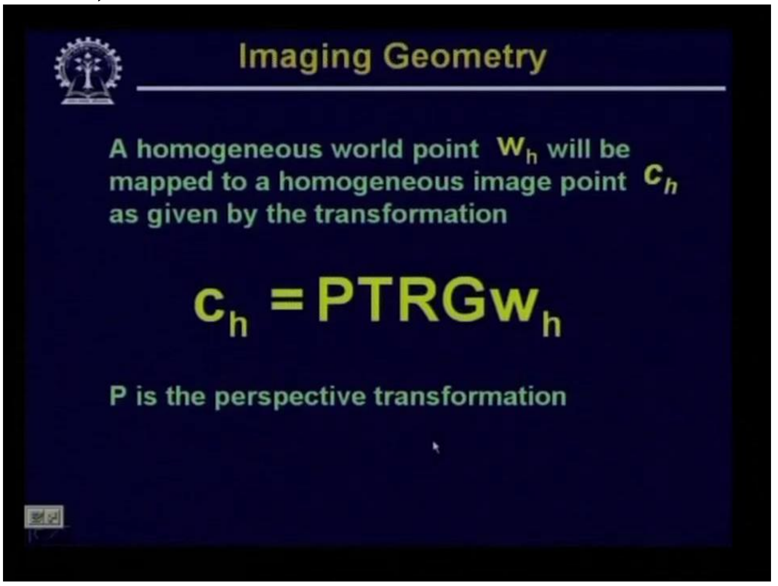
(Refer Slide Time 11:19)



x

0 1 0 minus r 2, 0 0 1 minus r 3, and then 0 0 0 1 So if apply all these transformations

(Refer Slide Time 11:30)

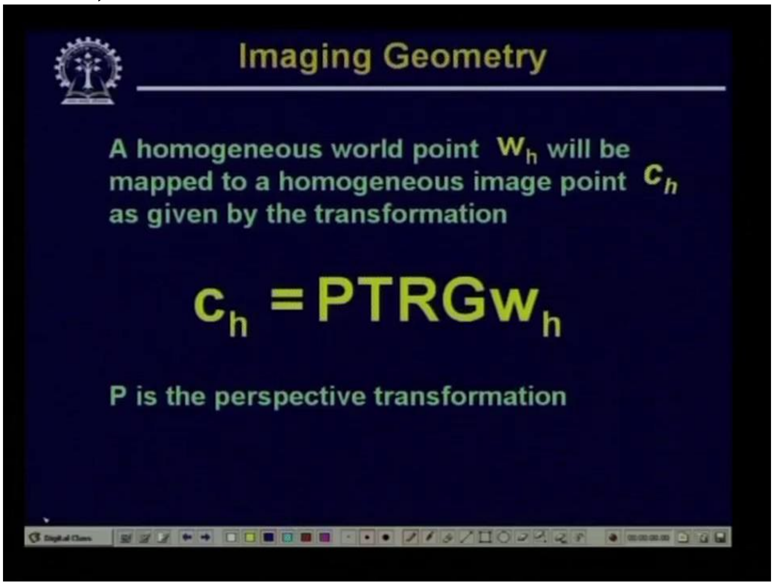


one after another and I represent the 3D world point w

(Refer Slide Time 11:36)

by the corresponding homogenous coordinates w h then you find

(Refer Slide Time 11:43)

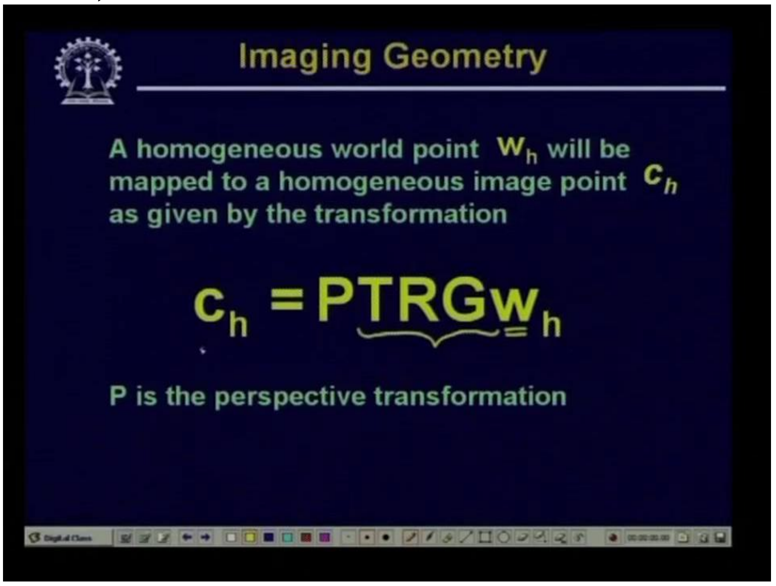


that these transformations t, r and g taken together on this homogenous coordinate w h gives you the homogenous transformed point w as seen by the camera. And once I have these transformed

(Refer Slide Time 12:08)

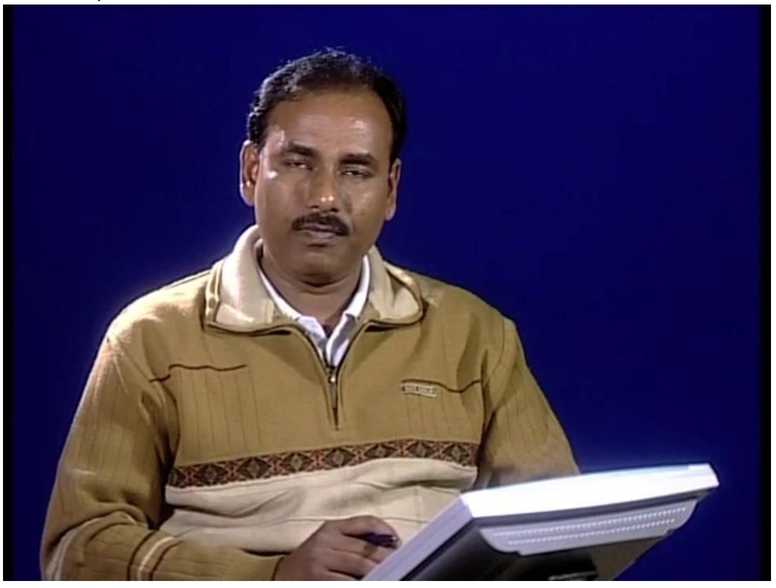
3D points then simply applying the perspective transformation on these transformed 3D points will give you the camera coordinate in the homogenous form. So now the camera coordinate

(Refer Slide Time 12:21)



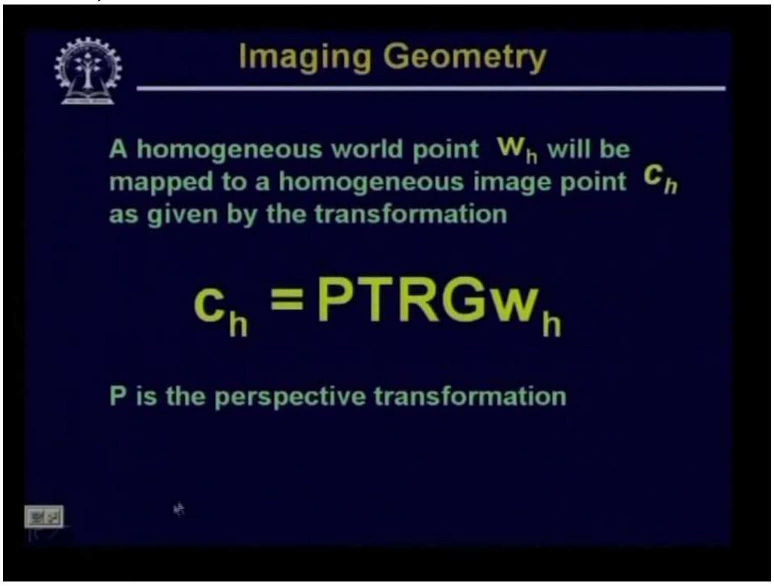
c h is given by p t r g into w h. So you remember that this coordinate comes in the homogenous form.

(Refer Slide Time 12:33)

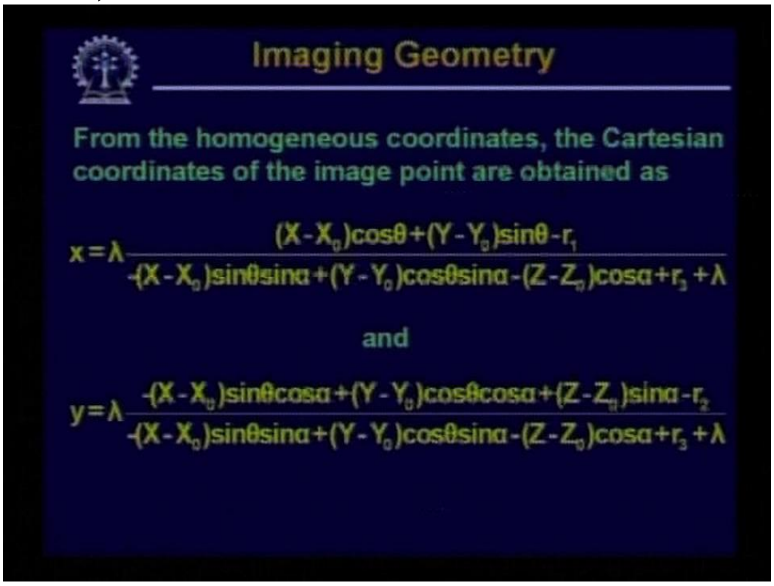


Then final operation that you have to do is to convert this homogenous coordinate c h into the corresponding Cartesian coordinate c so that Cartesian coordinate,

(Refer Slide Time 12:46)



(Refer Slide Time 12:47)



if I solve those equations, will come in this form, you can try to derive these equations that x equal to lambda into x minus x naught cosine theta plus y minus y naught sine theta minus r 1 divided by minus x minus y, x naught sine theta sine alpha plus y minus y naught cosine theta sine alpha minus z minus z naught cosine alpha plus r 3 plus lambda and the camera, image coordinate y is given by lambda, x minus x naught sine theta cosine alpha plus y minus y naught cosine theta cosine alpha plus z minus z naught sine alpha minus r 2 divided by x minus x naught sine theta sine alpha plus y minus y naught cosine theta sine alpha minus z minus z naught cosine alpha plus r 3 plus lambda.

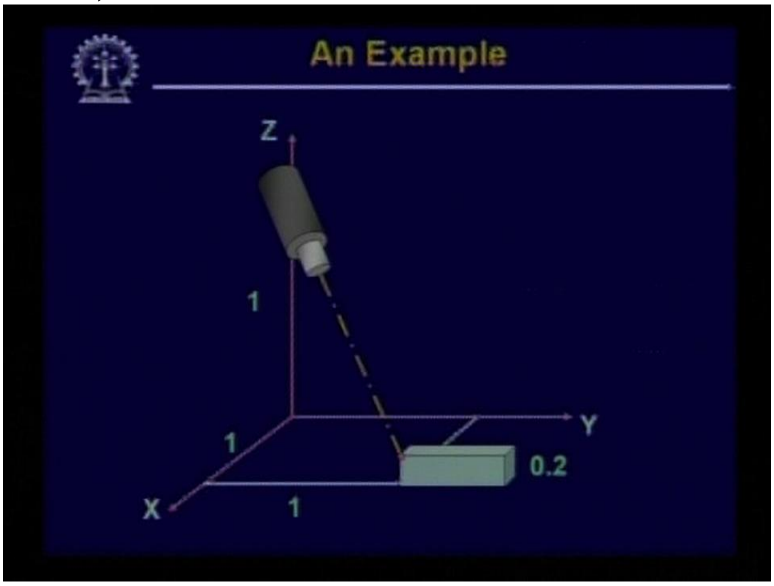
So these are

(Refer Slide Time 13:53)

the various transformation steps that we have to apply if I have a generalized imaging set up in which case the 3D coordinate axis and camera coordinate axis, they are not aligned. So the steps that we have to follow is first we assume that the camera coordinate axis and the 3D coordinate axis, they are perfectly aligned. Then give a set of transformations to the camera to bring to its given setup and apply the corresponding transformations but in the reverse direction to the 3D world coordinate points. So by applying these transformations to the 3D world coordinate points the image points, the 3D world coordinate points as seen by the camera will be obtained in the transformed point and after that if I apply the simple perspective transformation to this transformed 3D points, what I get is the image point corresponding to those transformed 3D world points.

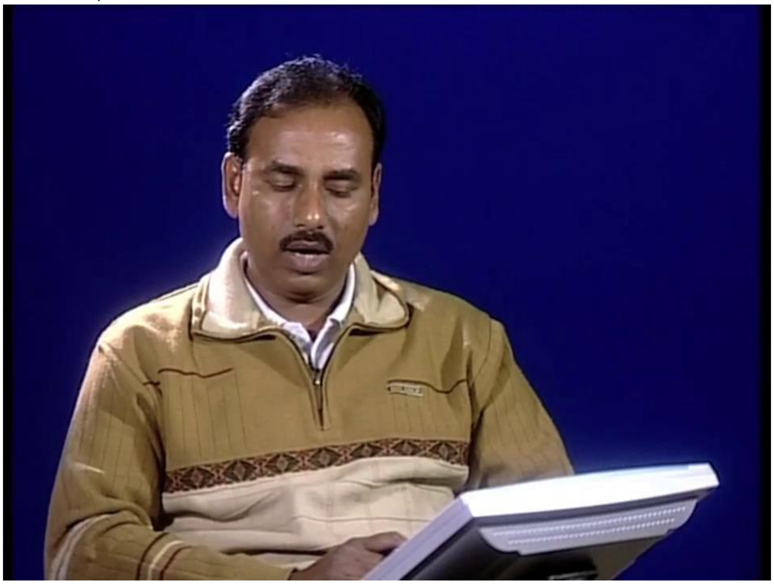
Now let us try to see an example to illustrate this operation.

(Refer Slide Time 15:02)



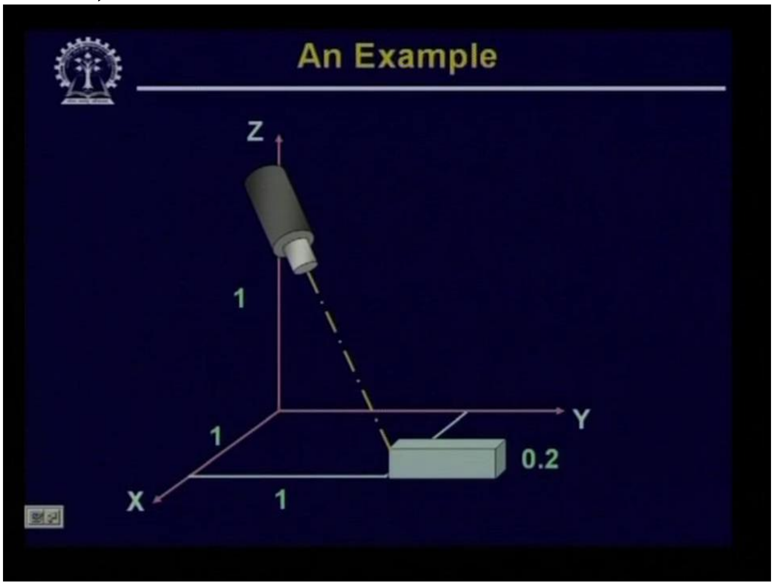
So let us take a figure where we assume that the camera or the center of the image plane, imaging plane of the camera is located at location 0 0 1 with respect to the 3D world coordinate system x y z. And we have an object placed in the x y plane where one of the corners of the object is at location 1 1 0 2, 0.2 and we want to find out that what will be the image coordinate for this particular 3D world point which is now a corner of this object

(Refer Slide Time 15:51)



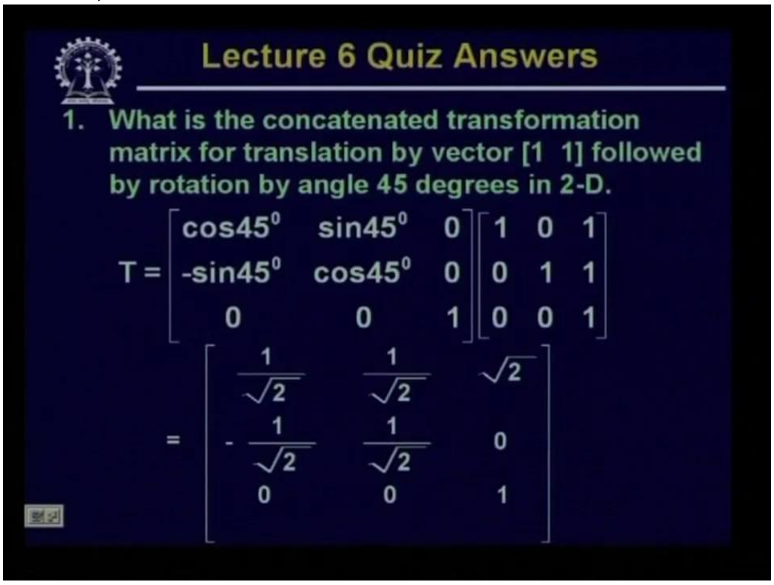
as placed in this figure. So what we will try to do is, we will try to apply the set of transformations to the camera plane one after another and try to find out

(Refer Slide Time 16:04)



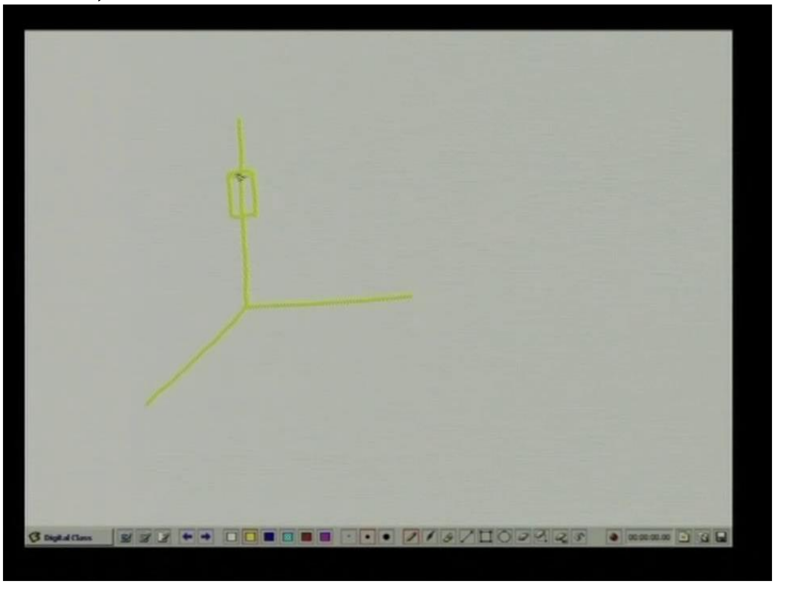
that what are the different transformations that we have to apply or what are the different corresponding transformations to the 3D world point that will bring, that will give us the world coordinate points, world points as seen by the camera. So initially I

(Refer Slide Time 16:24)



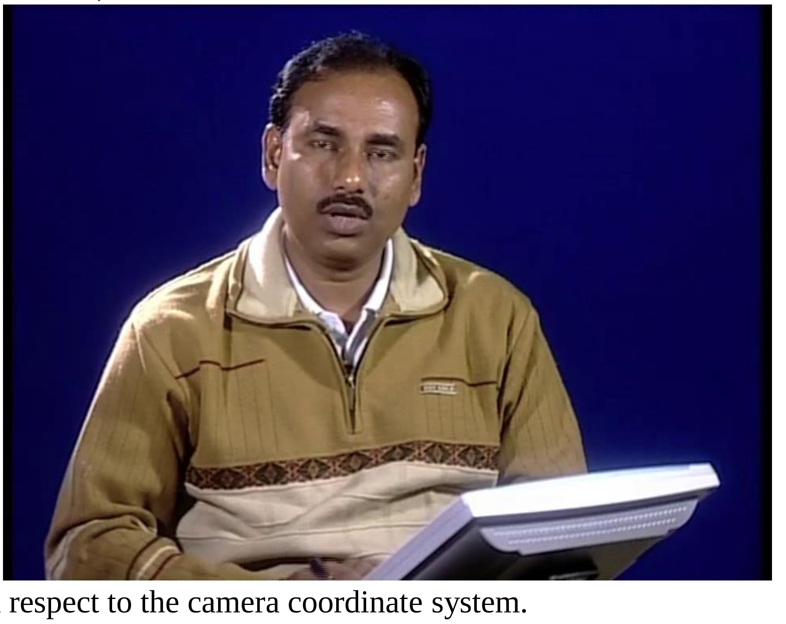
assume again that all the points or the image coordinate system and the camera coordinate system, they are perfectly aligned. Now after this assumption, what I have to do is, I have to give a displacement to the camera so, by the vector 0 0 1 so what I will do is I will bring the camera to a location here. So this is my

(Refer Slide Time 16:58)



camera where the image plane center is at location 0 0 1 So this is my x axis, this is my y axis, this is the z axis. Now if I do this transformation then you find that all the 3D points will be transformed by the vector

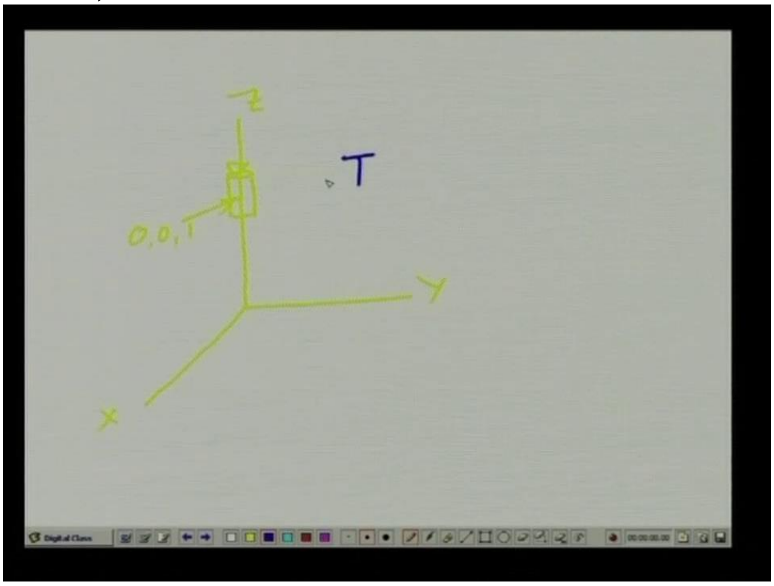
(Refer Slide Time 17:25)



0 0 minus 1 with respect to the camera coordinate system.

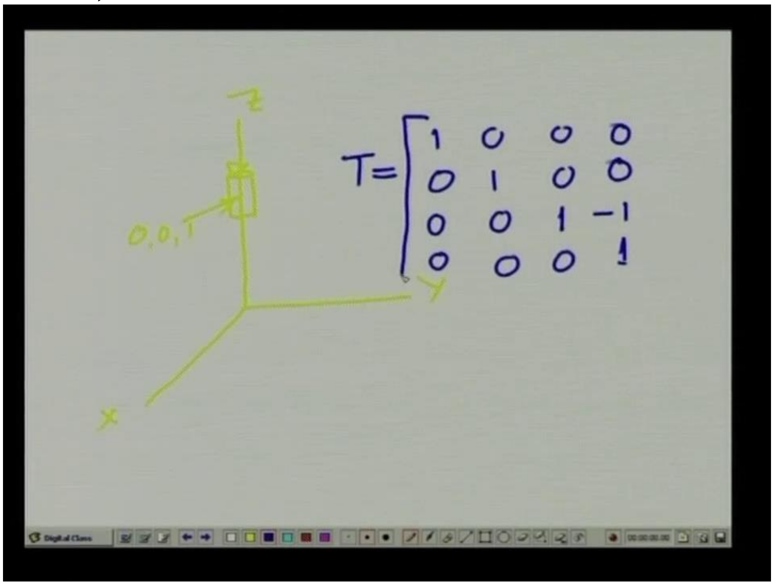
So the first transformation matrix which has to be applied to all the

(Refer Slide Time 17:35)



points in the 3D coordinate system is given by 1 0 0 0, 0 1 0 0, 0 0 1 minus 1 and then 0 0 0 1.

(Refer Slide Time 17:57)

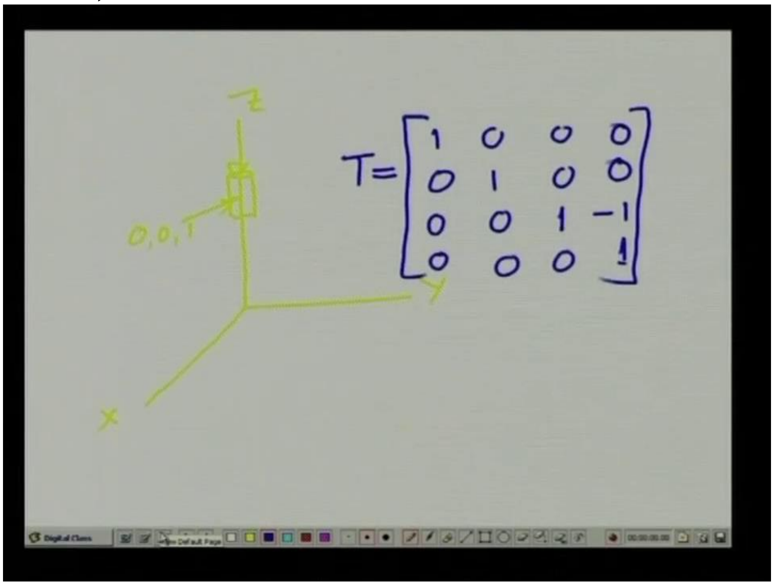


So this is the first transformation that has to be applied to all the 3D points Now after the camera is displaced by the vector 0 0 1, the next operation that you have to apply is to pan the camera by an angle 135 degree. I just forgot to mention that in that arrangement, that pan was 135 degree; the tilt was also 135 degree. So after the initial

(Refer Slide Time 18:28)

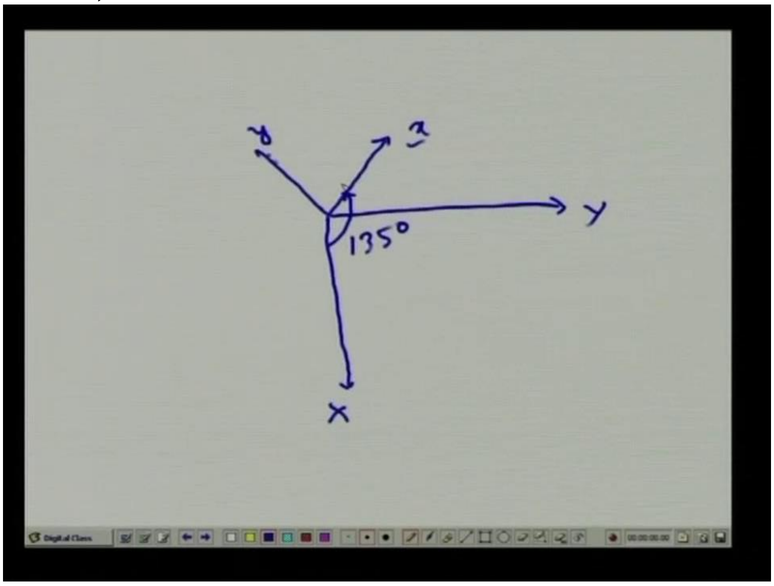
transformation, displacement of the camera by vector 0 0 1, we have to apply a pan of 135 degree to this camera. So if I

(Refer Slide Time 18:39)



represent that let us take a one, two-dimensional view. As we said, panning is nothing but rotation around z axis, so if I say that this is the x axis, this is the y axis then by panning we have to make an angle of 135 degree between the x axis of the camera coordinate system so the situation will be something like this. So this is the y axis of the camera coordinate system. This is the x axis of the camera coordinate system and by pan of 135 degree, we have to rotate the camera imaging plane in such a way that the angle between the x axis of the camera coordinate and the x axis of the 3D world coordinate is 135 degree.

(Refer Slide Time 19:38)

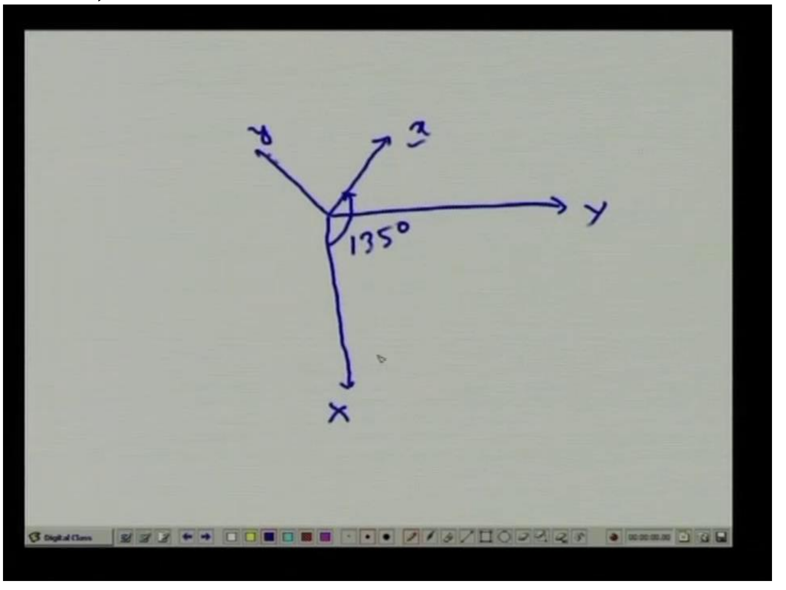


And once we do this, here you find that this rotation of the camera is in the anti-clockwise direction.

(Refer Slide Time 19:48)

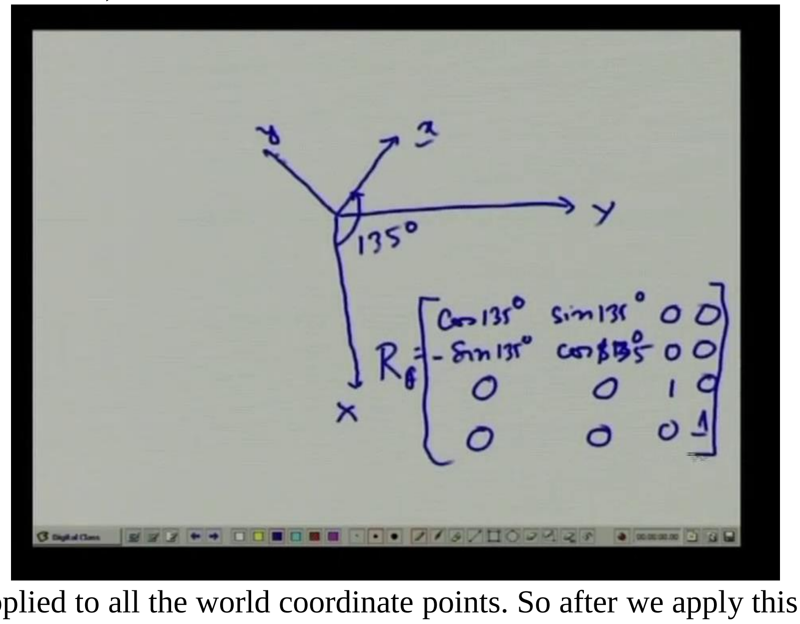
So the corresponding transformation on the 3D world points will be in the clockwise direction but by the same angle 135 degree. And the corresponding rotation matrix

(Refer Slide Time 19:59)



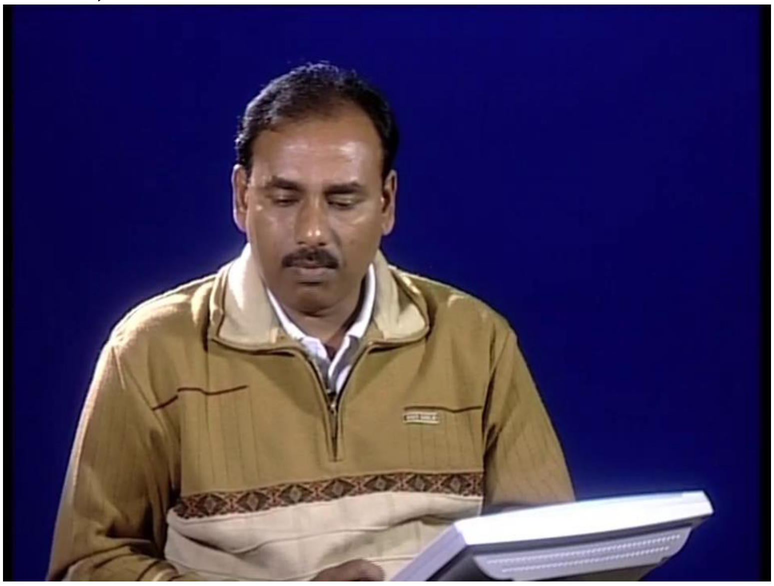
which is now given by r theta will be equal to cosine 135 degree sine 135 degree 0 0, then minus sine 135 degree cosine 135 degree 0 0, then 0 0 1 0, then 0 0 0 1. So this is the rotation transformation

(Refer Slide Time 20:35)



that has to be applied to all the world coordinate points. So after we apply this r theta the next operation that we have to perform is to tilt the camera by

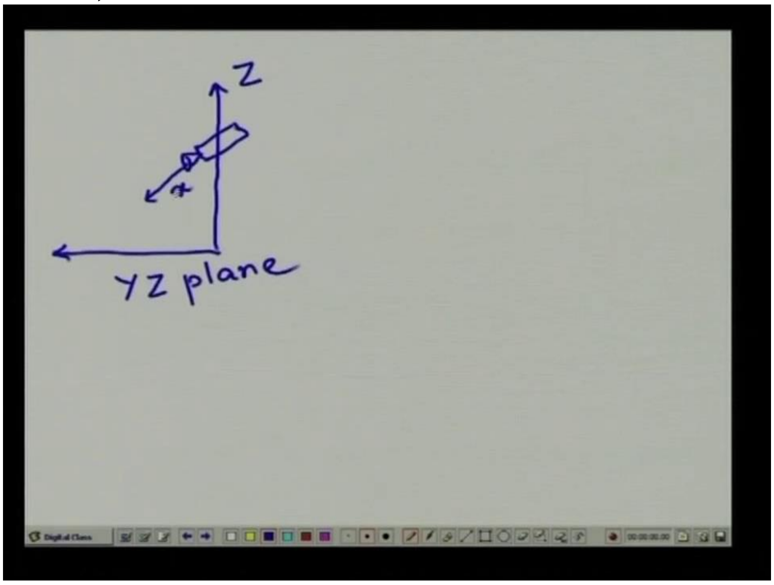
(Refer Slide Time 20:51)



an angle 135 degree

So again to have the look at this tilt operation, we take again a two dimensional view so the view will be something like this. So we take the z axis of the 3D world coordinate system and in this case it will be the y z plane of the 3D world coordinate system. And by tilt what we mean is something like this. This is the z axis of the camera coordinate system

(Refer Slide Time 21:30)

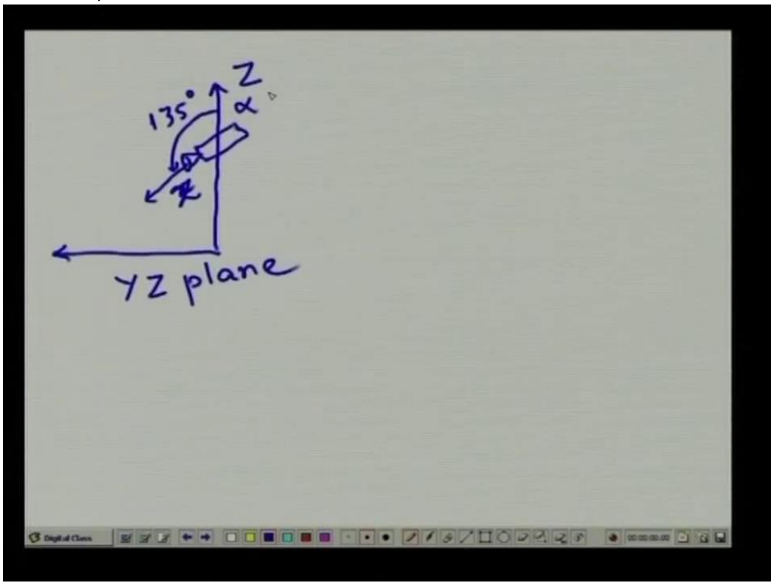


and the angle between the z axis of the 3D world coordinate system and the camera coordinate system is again 135 degree. So this is the angle, tilt angle alpha. So here again you find that the tilt is in the anti-clockwise direction so the corresponding transformation in the 3D world point will be

(Refer Slide Time 21:53)

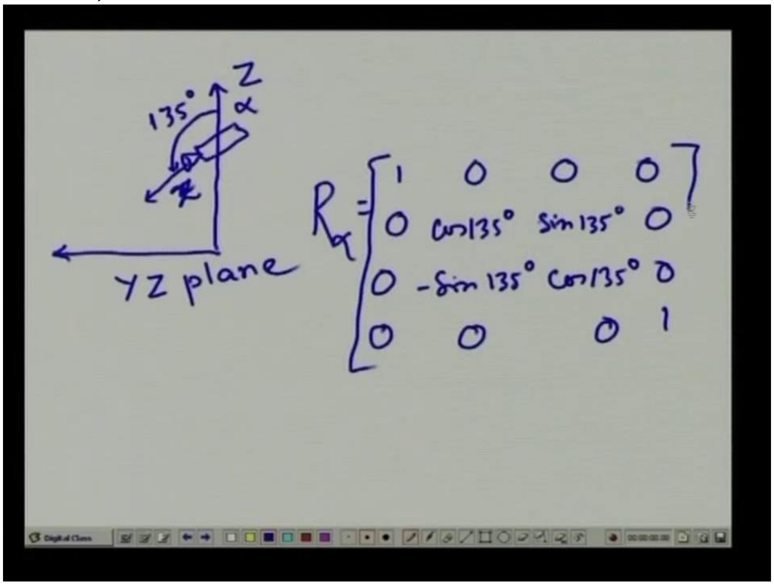
rotating the 3D world point by 135 degree in the clock wise direction around the

(Refer Slide Time 22:01)



x axis And the corresponding transformation matrix in this case will be r alpha is equal to 1 0 0 0, 0 cosine 135 degree sine 135 degree 0, 0 minus sine 135 degree cosine 135 degree 0, 0 0 0 1. So this is

(Refer Slide Time 22:37)

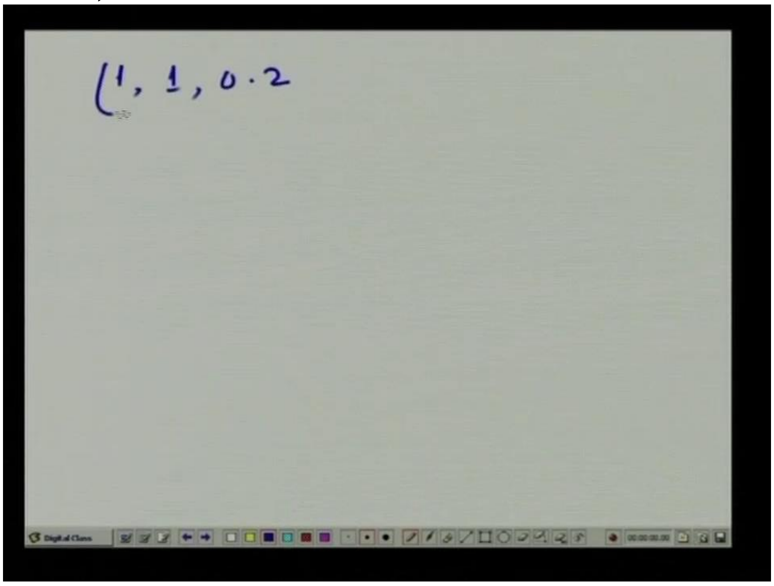


the transformation matrix that has to be applied to the tilt operation So after doing this, you find that the 3D world coordinate, the 3D world point for which we want

(Refer Slide Time 22:54)

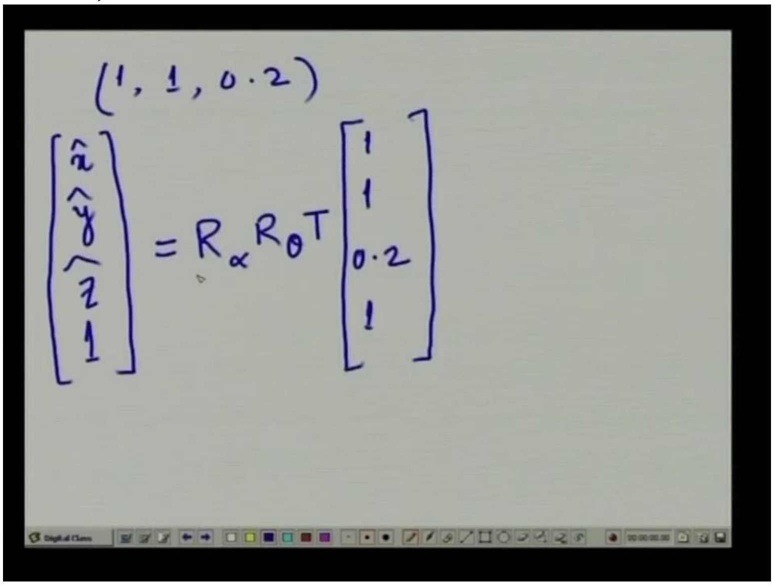
to find out the corresponding image point is given by 1 1 0 point 2.

(Refer Slide Time 23:03)



This is the 3D world coordinate point and after application of all these transformations, the transformed coordinate of this 3D world point, if we write it as x hat y hat z hat and this has to be represented in unified form, so this will be like this, it has to be r alpha r theta then t and the original world coordinate 1 1 0 point 2 and 1 in the unified form. Now if I compute this

(Refer Slide Time 23:46)

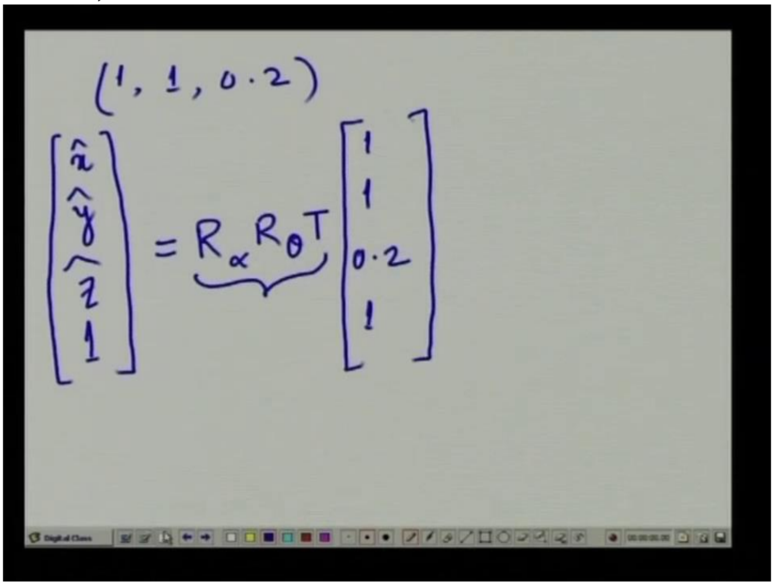


r alpha, r theta and t using the transformation matrix that

(Refer Slide Time 23:50)

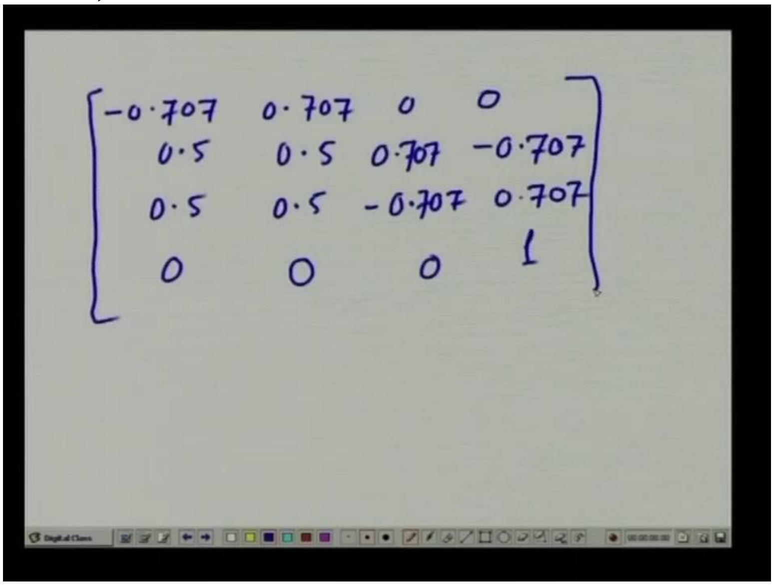
we have just computed, that we have just derived, you will find that this transformation matrix

(Refer Slide Time 23:57)



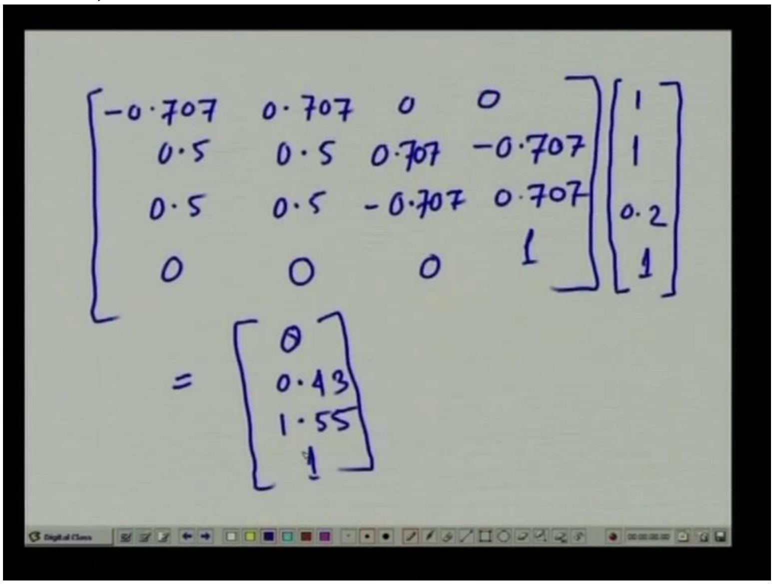
can be computed as minus 0 point 707 0 point 707 0 0, then 0 point 5 0 point 5 0 point 707 -0 point 707, then again 0 point 5 0 point 5 minus 0 point 707 0 point 707 then 0 0 0 1. This is the overall transformation matrix which takes care of

(Refer Slide Time 24:45)



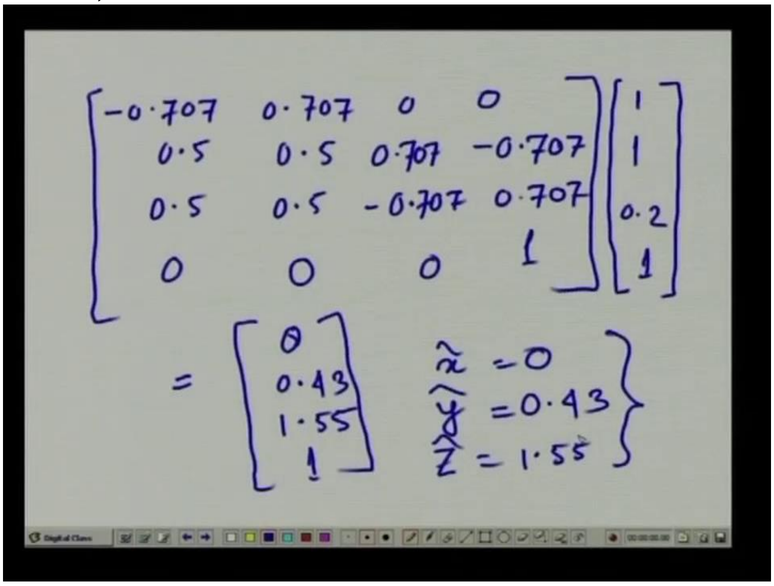
the transform translation of the image plane, then pan by angle theta and also tilt by angle alpha So if I apply this transformation to my original 3D world coordinates which was 1 1 0 point 2 then 1 then what I get is the coordinates of the point as observed by the camera. So if you compute this, you will find that this will come in the form 0 0 point 43 1 point 55 and then 1 again this is in the unified form. So the

(Refer Slide Time 25:27)



corresponding world the corresponding Cartesian coordinates will be given by x hat equal to 0, y hat equal to 0 point 43 and z hat equal to 1 point 55. So these are the coordinates for the same 3D point

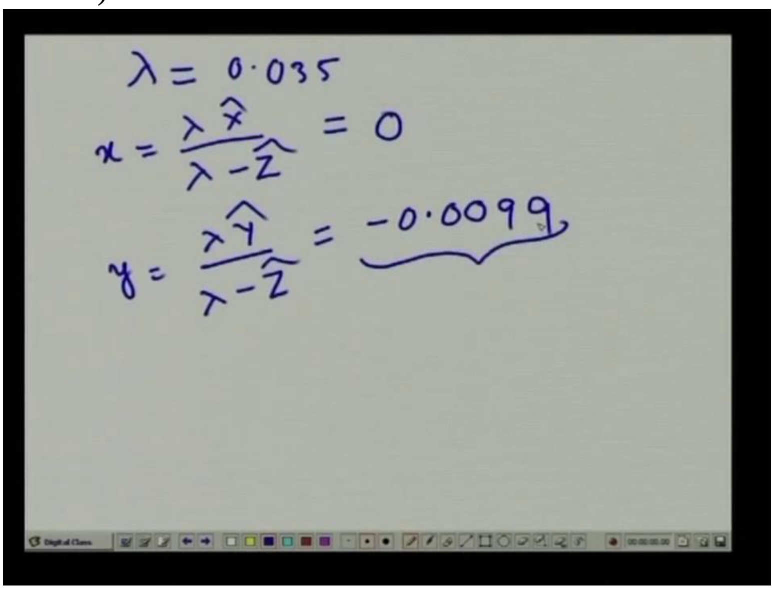
(Refer Slide Time 25:53)



as seen by the camera.

So now what we have to do is we have to apply the perspective transformation to this particular point. So if I apply the perspective transformation and if assume that the focal length of the camera is 0point 035 then we obtain the image coordinates as x equal to lambda x hat divided by lambda minus z hat which will be in this case, of cause 0 and y equal to lambda y hat divided by lambda minus z hat, which if you compute, this will come as minus 0 point 0099. So these are the image coordinates of the world coordinate point that we have considered. Now note that the y coordinate in the image plane

(Refer Slide Time 26:54)

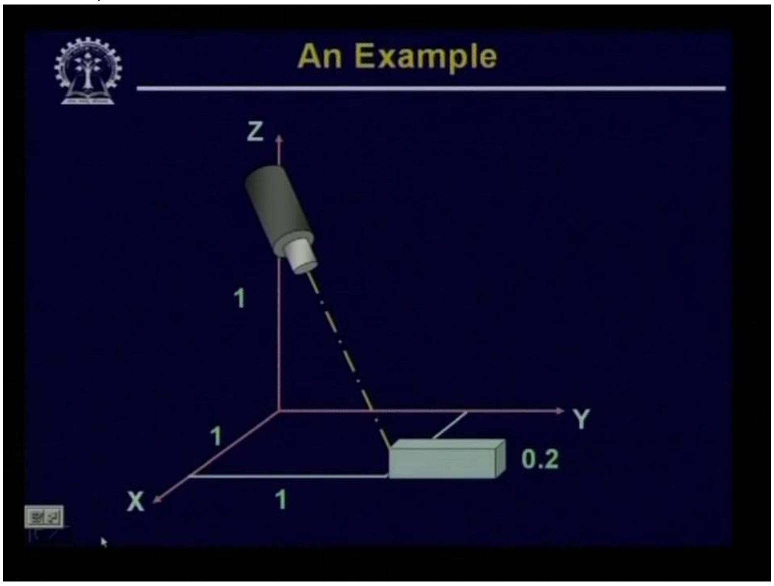


has come out to be negative This is obvious because the original 3D world coordinate as obtained by, after applying the transformations came out to be positive so obviously in case of image plane, there will be an inversion so this value of y coordinate will come out to be negative. So this particular example

(Refer Slide Time 27:17)

illustrates the set of transformations that we have to apply followed by the perspective transformation so that we can get the image point for any arbitrary point in the in the 3D world.

(Refer Slide Time 27:35)



So with this we complete our discussion on the different transformations and the different imaging models that

(Refer Slide Time 27:46)

we have taken