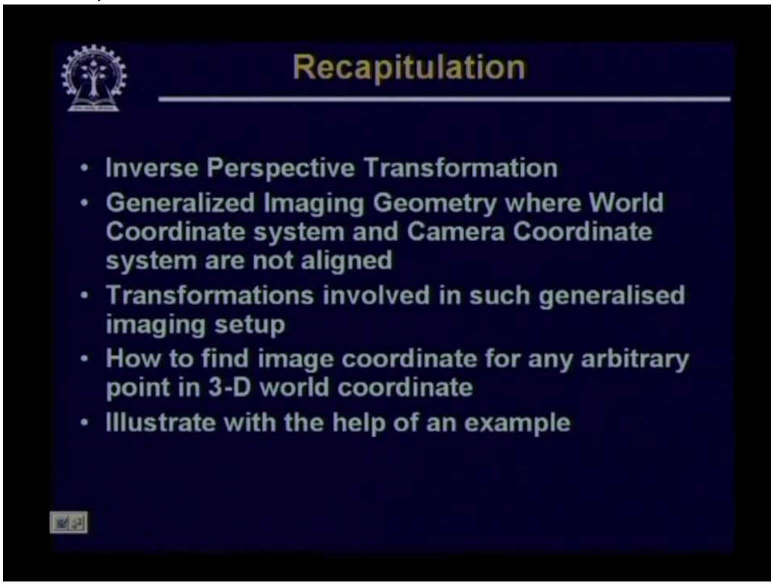
Digital Image Processing
Prof. P. K. Biswas
Department of Electronics and Electrical Communications Engineering
Indian Institute of Technology, Kharagpur
Module 03 Lecture Number 14
Image Geometry - 2

(Refer Slide Time 00:20)

Hello, welcome to the video lecture series on Digital Image Processing. In a 3D coordinate system where the coordinate system and the camera coordinate system are not perfectly aligned, in that case what are the set of transformations which are to be applied to the points in the 3D world coordinate system which will be transformed in the form as seen by a camera then followed by that if we apply the perspective transformation then we get the image coordinates for different points in the 3D world coordinate system.

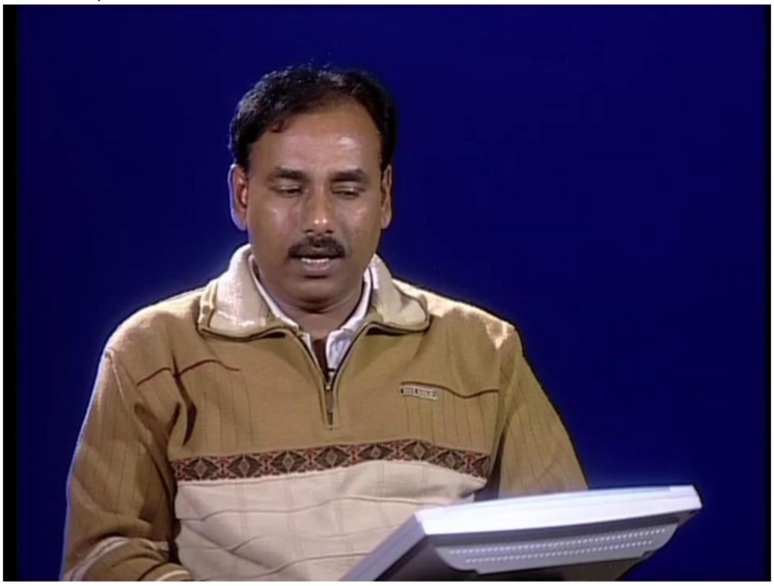
So what we have seen in the last class is

(Refer Slide Time 01:02)



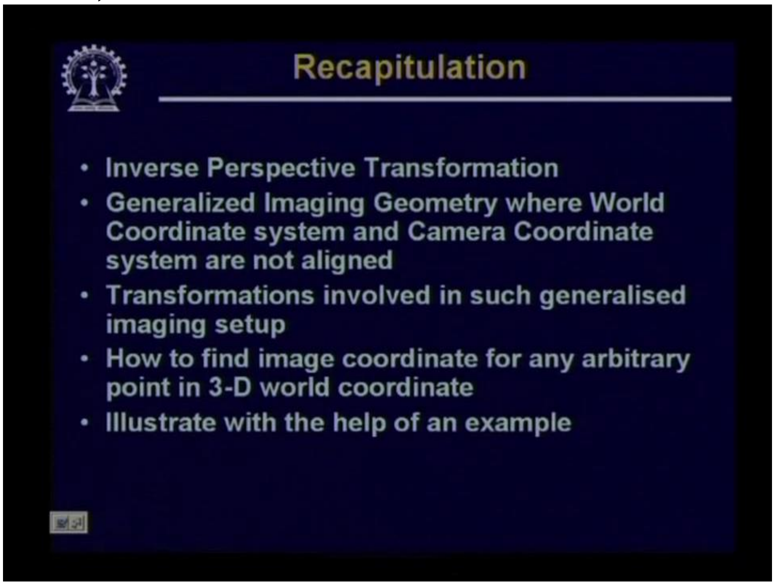
once we have the image points in an image plane,

(Refer Slide Time 01:07)



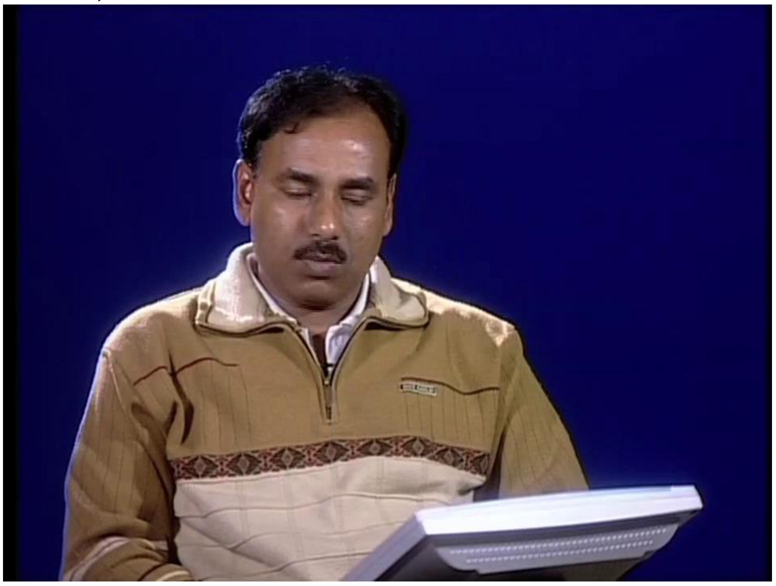
how to apply the inverse perspective transformation to get the equation of the straight line so that the points on that straight line map to a particular image point on the imaging plane Then we have seen a generalized imaging geometry

(Refer Slide Time 01:24)



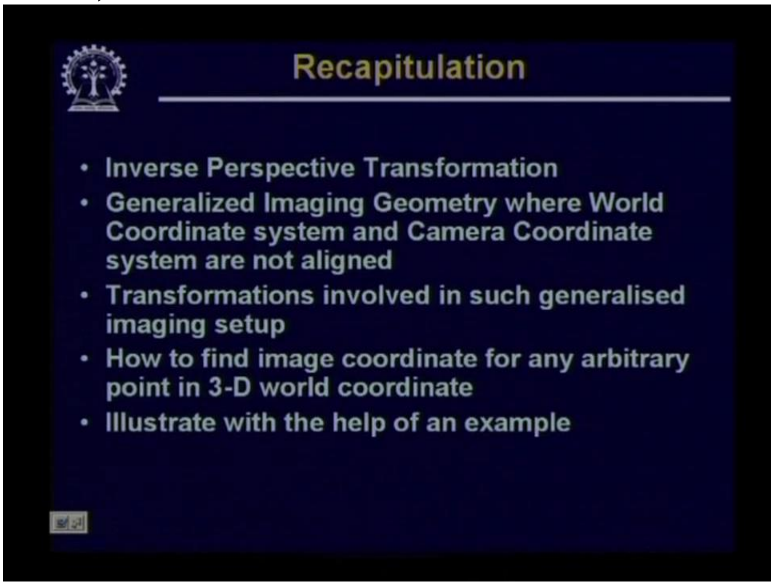
where the world coordinate system and the camera coordinate system are not aligned and we have also discussed the set of transformations

(Refer Slide Time 01:34)



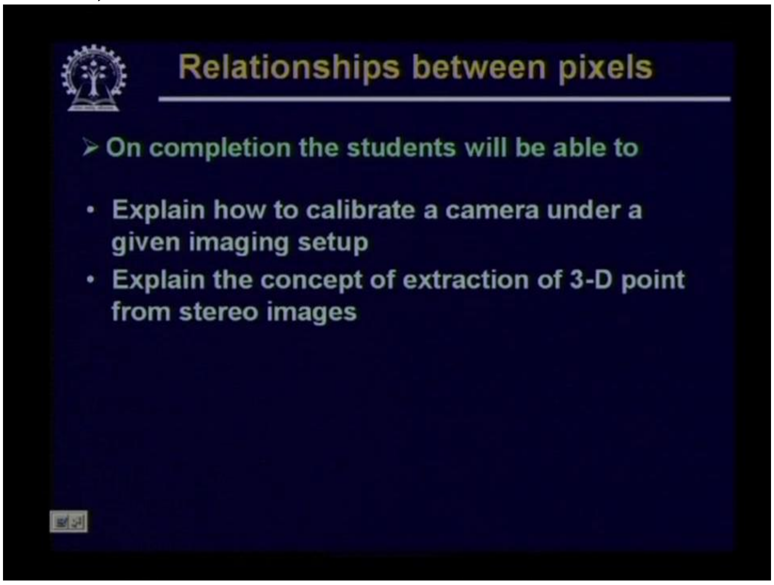
which are involved in such generalized imaging setup and then

(Refer Slide Time 01:39)



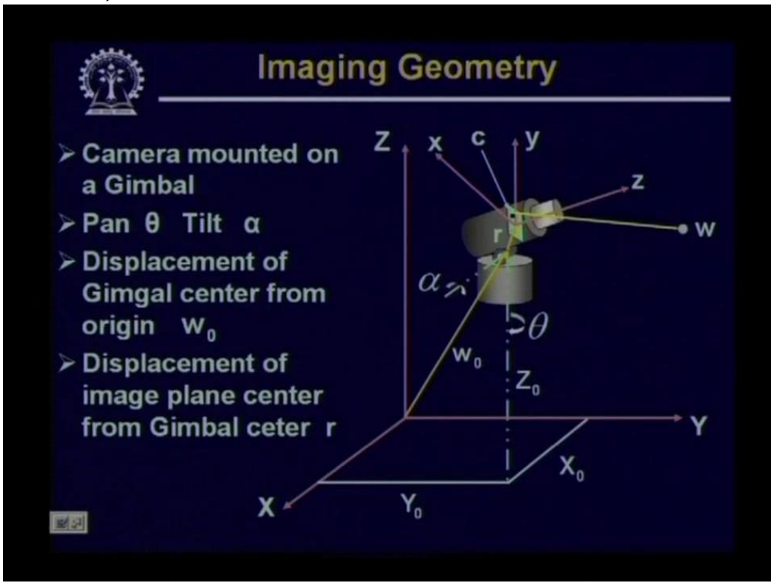
we have also seen how to find image coordinate for any arbitrary point in the 3D world coordinate system in such a generalized imaging setup and the concept we have illustrated with the help of an example

(Refer Slide Time 01:58)



In today's lecture we will see that given an imaging setup, how to calibrate the camera and then we will also explain the concept of how to extract the 3D point from two images which is also known as stereo images.

(Refer Slide Time 02:19)

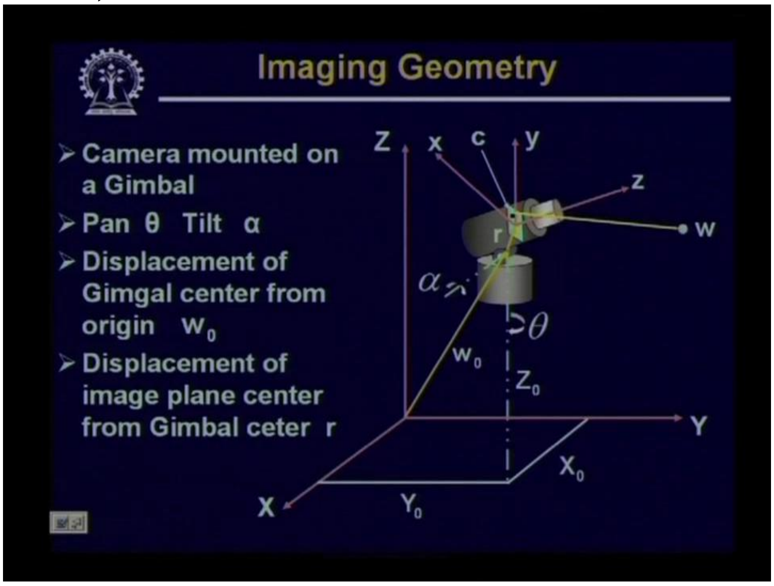


So in the last class that we have done is we have given an imaging setup like this while the

(Refer Slide Time 02:27)

3D world coordinate system is given by capital X capital Y capital Z. In this world coordinate system we had placed a camera while that

(Refer Slide Time 02:37)

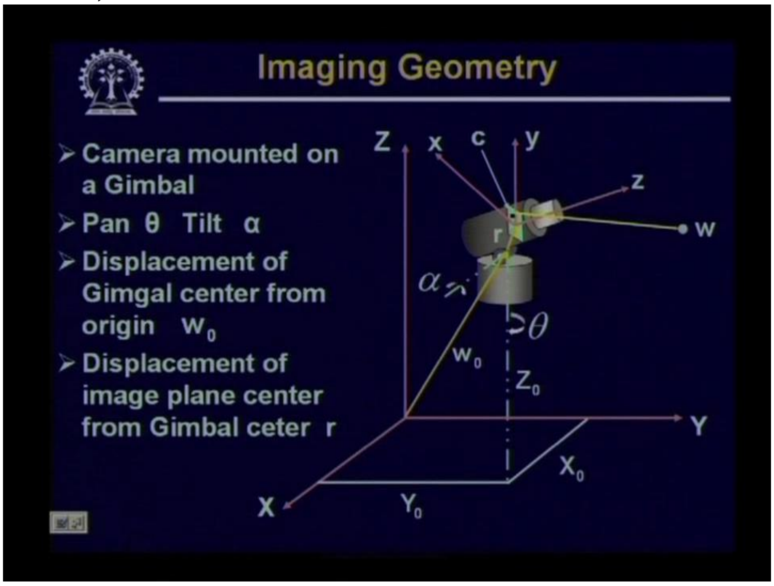


camera coordinate system is given by small x small y small z and we had assumed that camera is placed, is mounted on a gimbal where the gimbal is displaced from the origin of the world coordinate system

(Refer Slide Time 02:54)

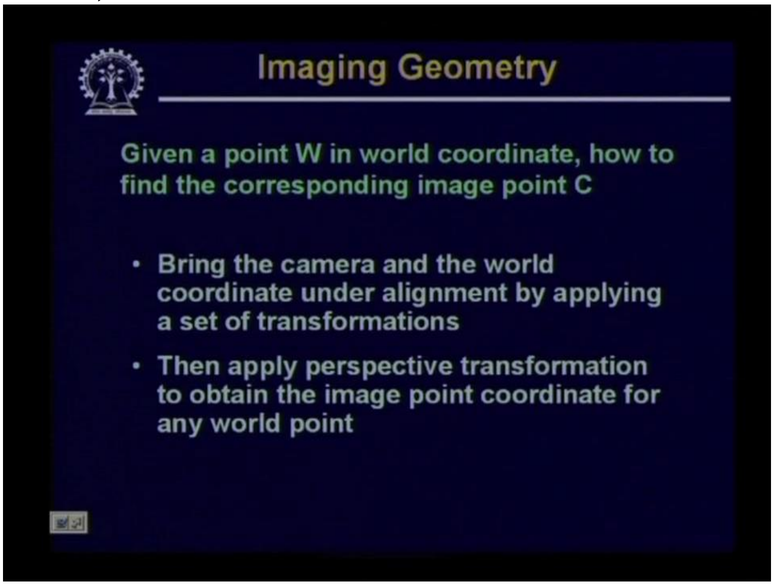
by a vector w naught and the center of the camera

(Refer Slide Time 02:58)



is displaced from the gimbal by a vector r, the camera is given a pan of angle theta and it is also given a tilt of angle alpha and in such a situation if w is a point in the 3D world coordinate system we have seen that how to find out the corresponding image point corresponding to point w in the image plane of the camera. So for that we have done

(Refer Slide Time 03:31)

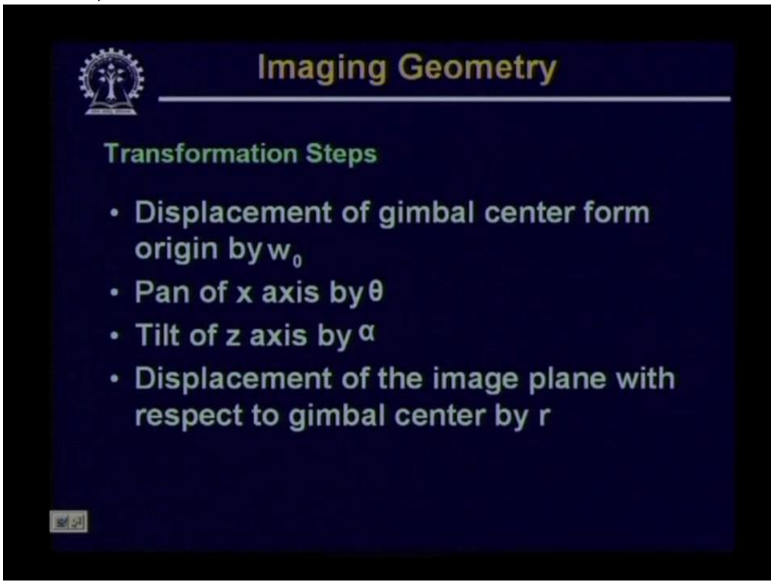


a set of transformations and with the help of the set of transformations what we have done is

(Refer Slide Time 03:37)

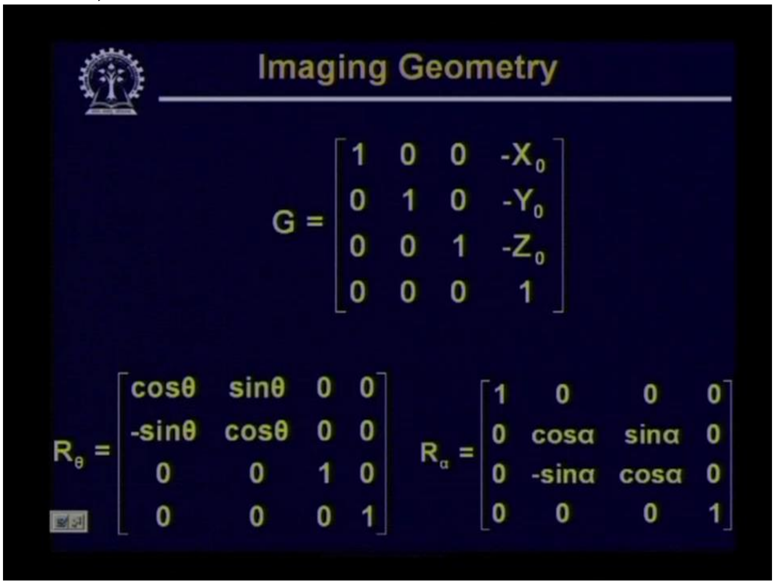
we have brought the 3D world coordinate system and camera coordinate system in alignment and after the 3D world coordinate system and the camera coordinate system are perfectly aligned with that set of transformations then we have seen that we can find out the image point corresponding to any 3D world point by applying the perspective transformation. So the type of transformations that

(Refer Slide Time 04:06)



we have to apply is first we have to apply that transformation for the displacement of the gimbal center from the origin of the 3D world coordinate system by vector w naught followed by a transformation corresponding to the pan of x axis of the camera coordinate system by theta which is to be followed by a transformation corresponding to a tilt of the z axis of the camera coordinate system by angle alpha and finally the displacement of the camera image plane with respect to gimbal center by vector r. So

(Refer Slide Time 04:48)

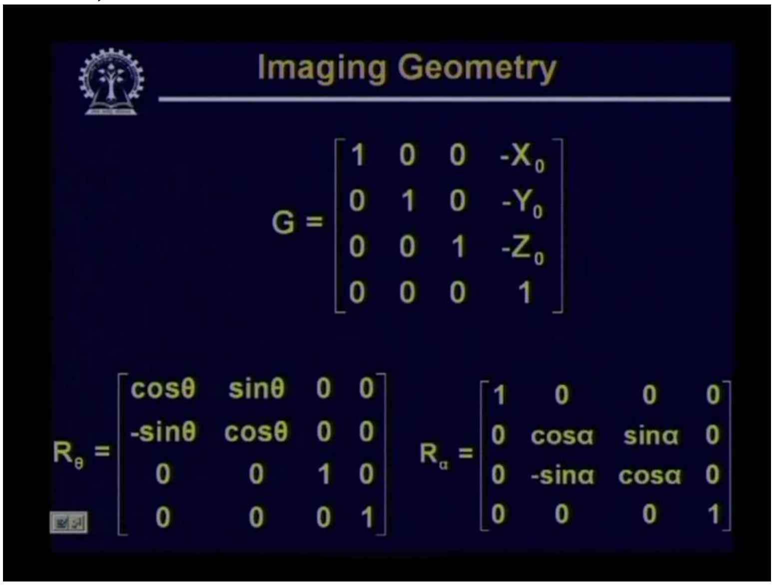


the transformation, the first transformation which translates the gimbal center

(Refer Slide Time 04:53)

from the world, origin of the world coordinate system

(Refer Slide Time 04:57)

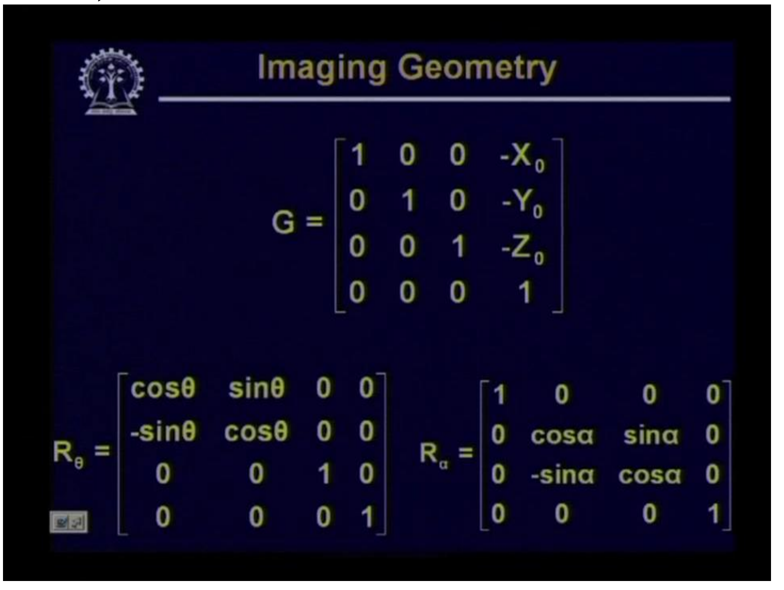


by vector w naught is given by the transformation matrix g which is in this case 1 0 0 minus x naught, 0 1 0 minus y naught, 0 0 1 minus z naught and 0 0 0 1. The pan of the x axis of the

(Refer Slide Time 05:18)

camera coordinate system by an angle theta is given by the transformation matrix r theta where r theta in this case is cosine theta sine theta 0 0, then

(Refer Slide Time 05:30)

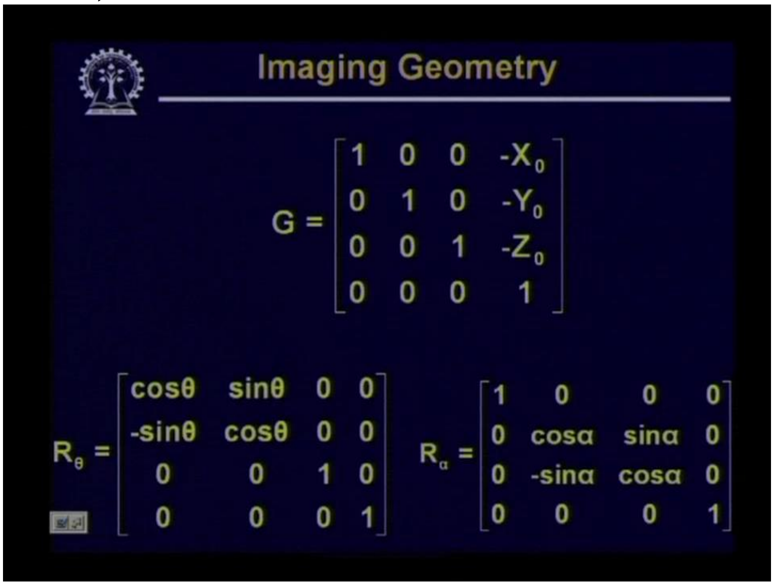


minus sin theta cosine theta 0 0, then 0 0 1 0 and 0 0 0 1 Similarly the tilt, the transformation matrix corresponding to the tilt by

(Refer Slide Time 05:45)

angle alpha is given by the other transformation matrix r alpha which

(Refer Slide Time 05:51)



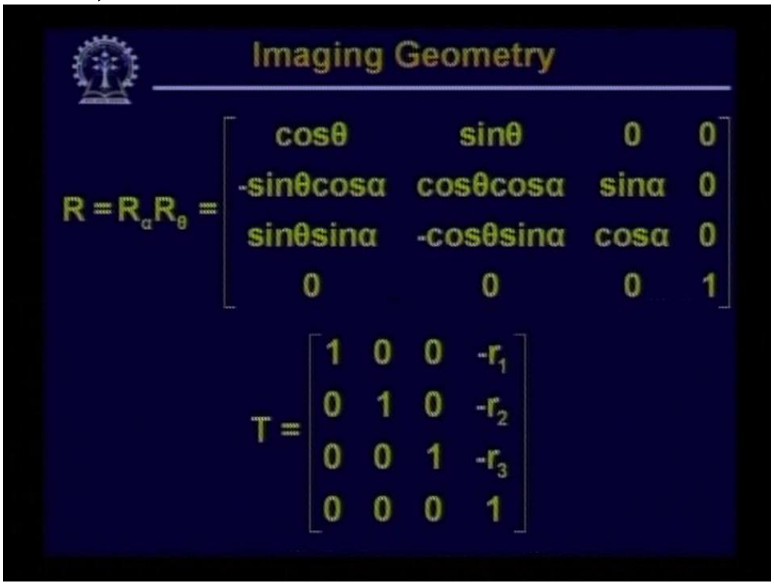
in this case is 1 0 0 0, 0 cosine alpha sine alpha 0, then 0 minus sine alpha cosine alpha 0 and then then 0 0 0 1.

Then the next transformation we have to apply is the transformation of the

(Refer Slide Time 06:15)

center of the image plane with respect to gimbal center by the vector r So if we assume that r has the components, the vector r has components r 1, r 2 and r 3 in x y along the x direction, y direction and z direction of the 3D world coordinate system then corresponding transformation matrix with respect to this translation is given by t equal to

(Refer Slide Time 06:43)

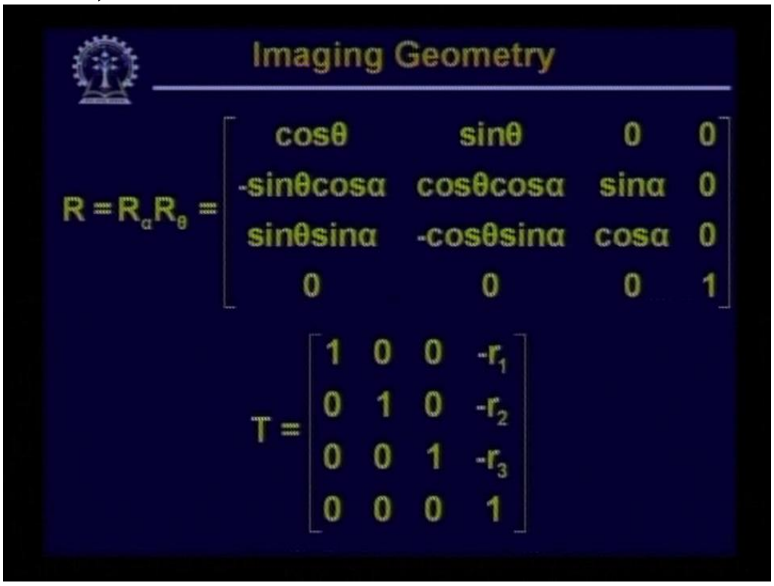


1 0 0 minus r 1, 0 1 0 minus r 2, 0 0 1 minus r 3, and then 0 0 0 1 We have also seen in the last class that the rotation matrices

(Refer Slide Time 07:00)

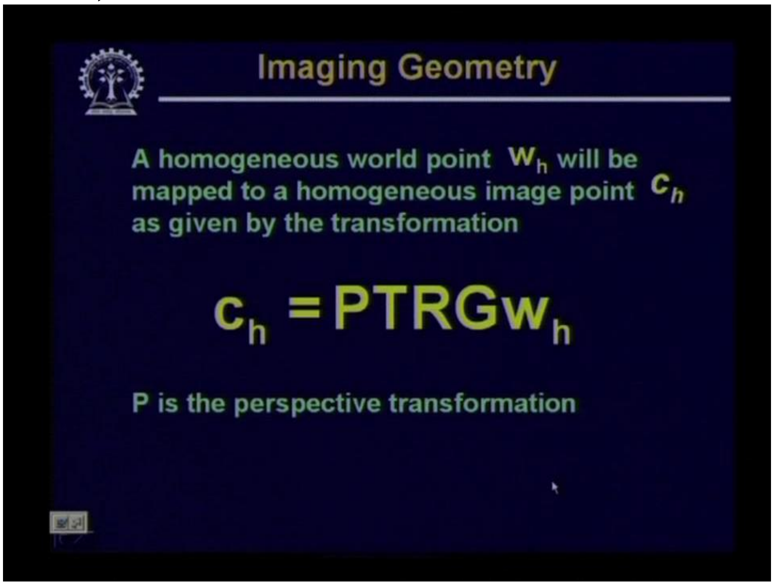
r theta and r alpha can be combined together to give a single transformation matrix r which is nothing but the product of r alpha and r theta. Once we get these

(Refer Slide Time 07:14)



transformation matrices

(Refer Slide Time 07:16)

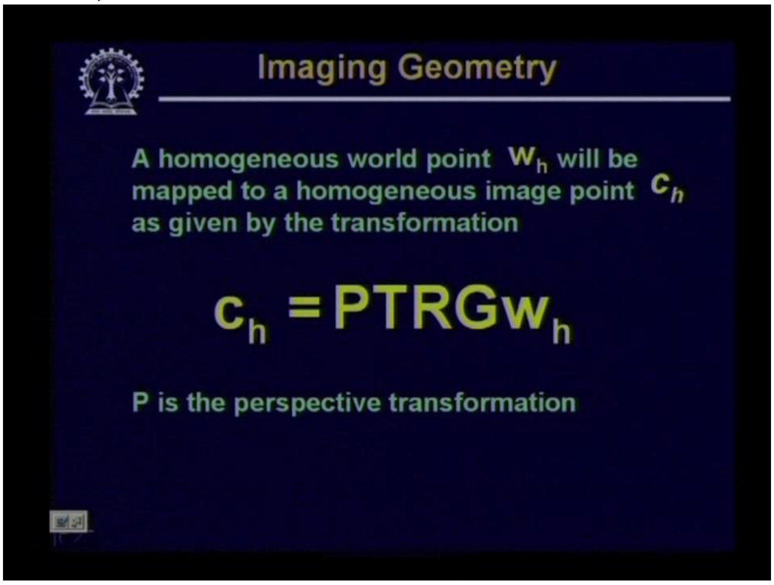


then after the transformation first by the

(Refer Slide Time 07:21)

translation matrix g then by the rotation matrix r followed by the second, second translation matrix what we do is we align the coordinate system of the camera with the 3D world coordinate system. That means now that every point in the 3D world will have a transformed coordinate as seen by the camera coordinate system. So once we do this then finally applying the perspective transformation to these

(Refer Slide Time 07:51)

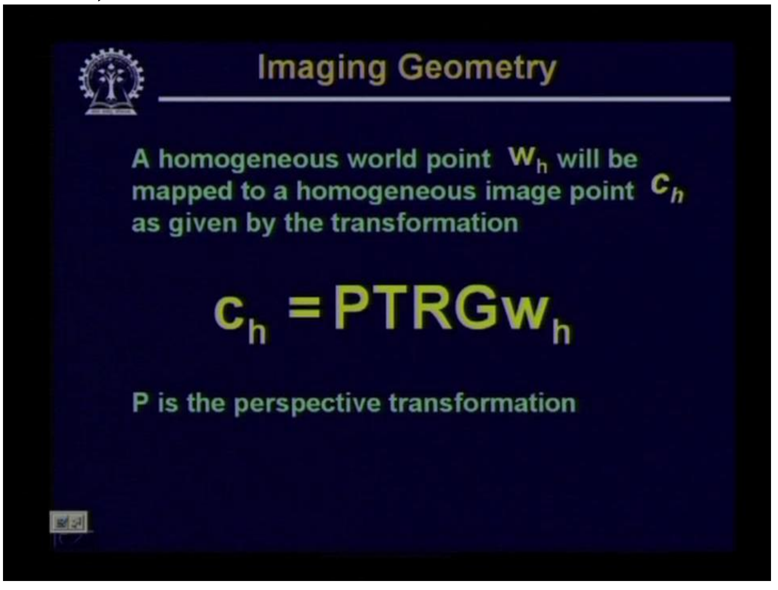


three-dimensional world coordinate systems gives us the coordinate

(Refer Slide Time 07:56)

of the point in the image plane for any point in the 3D world coordinate system So here you find the

(Refer Slide Time 08:04)

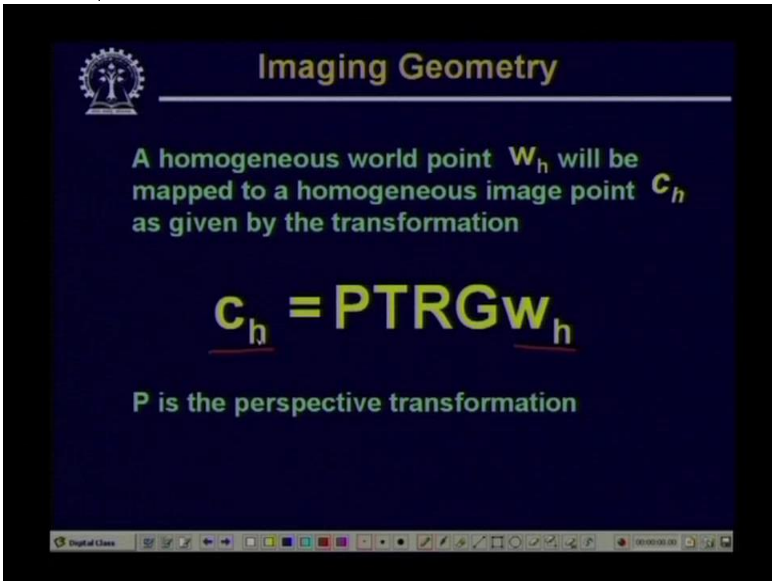


final form of the expression is like this, that both of the world coordinate system and the camera coordinate system, in this case they are represented in the homogenous form. So w h is the homogenous coordinate corresponding to the world coordinate w and c h is the homogenous form of the image coordinate c. So for a world point w whose homogenous coordinate is

(Refer Slide Time 08:35)

represented by w h here you can find that I can find out the image coordinate of the point w again in the homogenous form which is given by this matrix equation that c h is equal to

(Refer Slide Time 08:52)

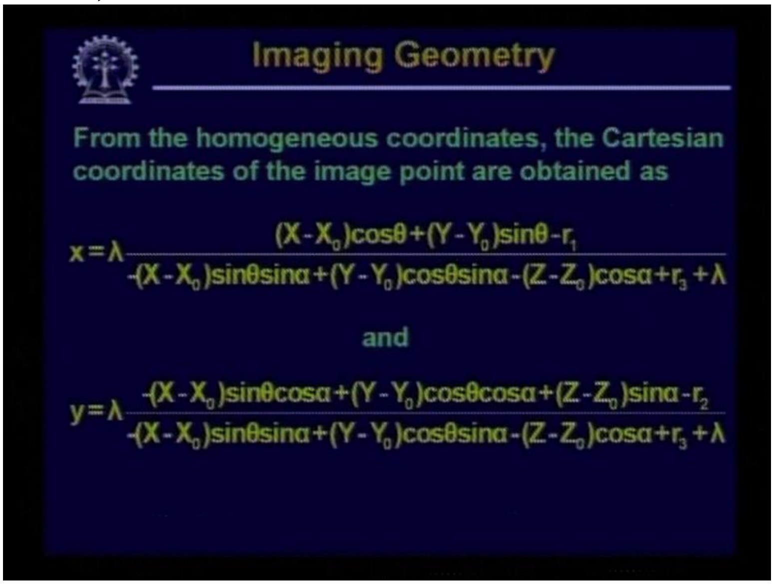


p t r g and w h And here you note that each of these transformation matrices that is p, t, r and g. all of them, they are of dimension

(Refer Slide Time 09:05)

4 by 4 So when I multiply all these matrices together to give a single transformation matrix, then the dimension of that transformation matrix will also 4 by 4. So what we have now is of this form. So after doing this

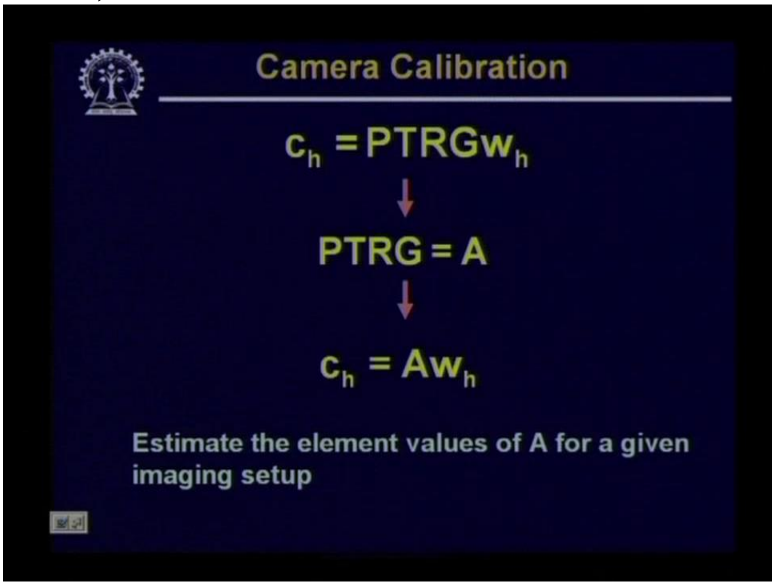
(Refer Slide Time 09:24)



transformation I can find out the image coordinate of the corresponding point w where x and y coordinates will be given by these expressions.

So after doing this what

(Refer Slide Time 09:38)

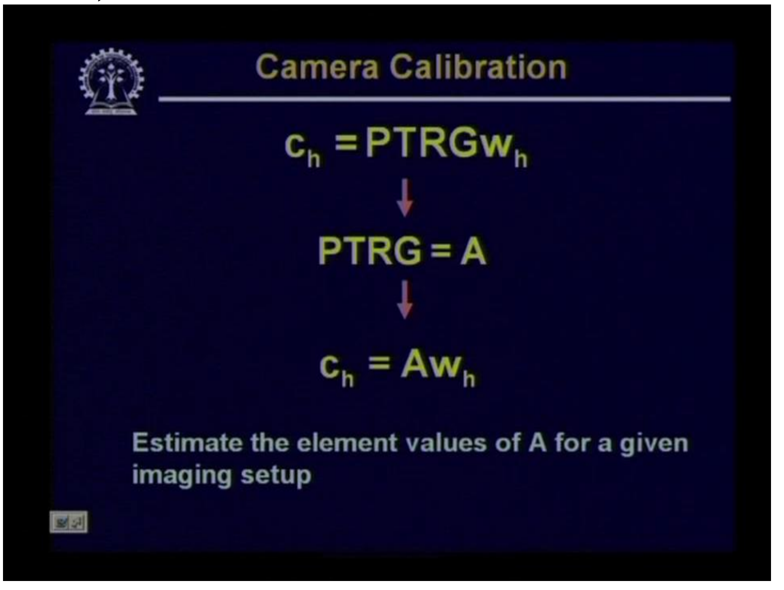


I have is I have an equation, a transformation or a matrix equation so that for any world point w I can find out what is the homogenous coordinate

(Refer Slide Time 09:54)

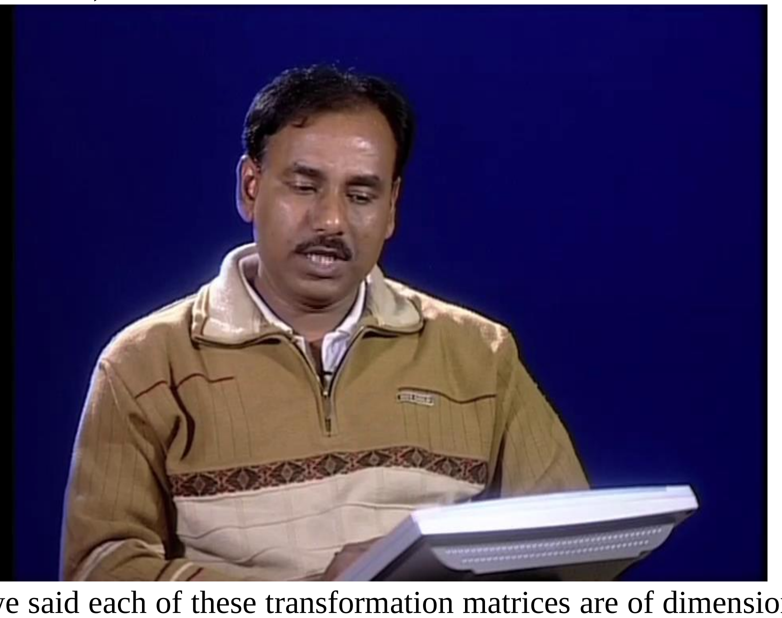
of the image point corresponding

(Refer Slide Time 09:57)



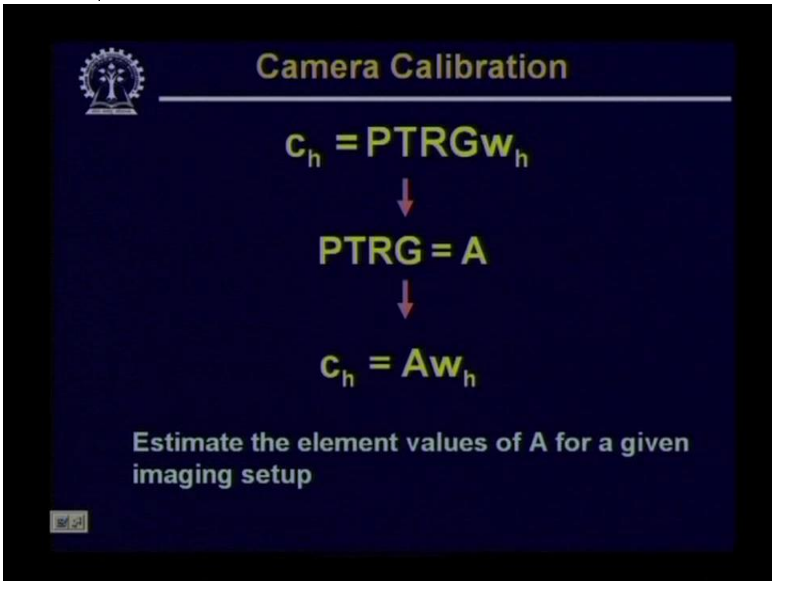
to that particular point w Now the transformation which is involved that is

(Refer Slide Time 10:03)



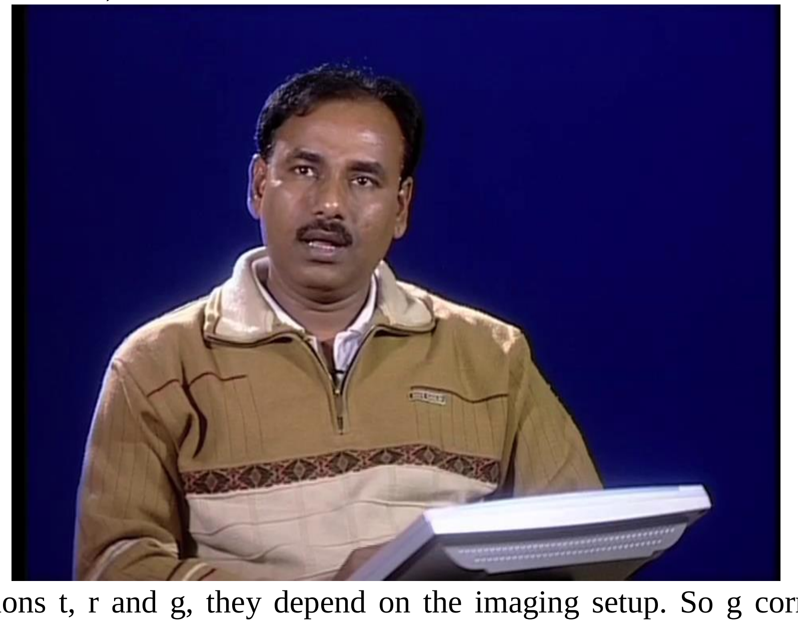
p, t, r and g, as we said each of these transformation matrices are of dimension 4 by 4. So the combined transformation matrix, if I represent this by a matrix a then this matrix a will also be a 4 by 4 matrix and now the inter-transformation equation in matrix form will be c h is equal to

(Refer Slide Time 10:29)



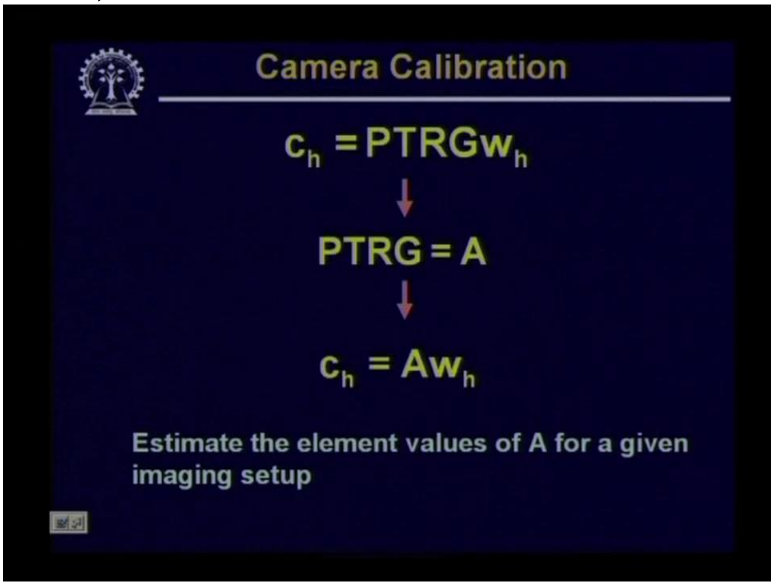
a into w h Now find that given a particular setup

(Refer Slide Time 10:36)



the transformations t, r and g, they depend on the imaging setup. So g corresponds to the transformation of the gimbal center from the origin of the 3D world coordinate system, r corresponds to the pan angle and the tilt angle and t corresponds to translation of the image plane center from the gimbal center. So these 3 transformation matrices depend upon the geometry of the imaging system where as the other transformation matrix that is p or perspective transformation matrix, this is entirely a property of the camera because you will find the components of this transformation matrix p has a term lambda which is equal to the wavelength the focal length of the camera.

(Refer Slide Time 11:26)



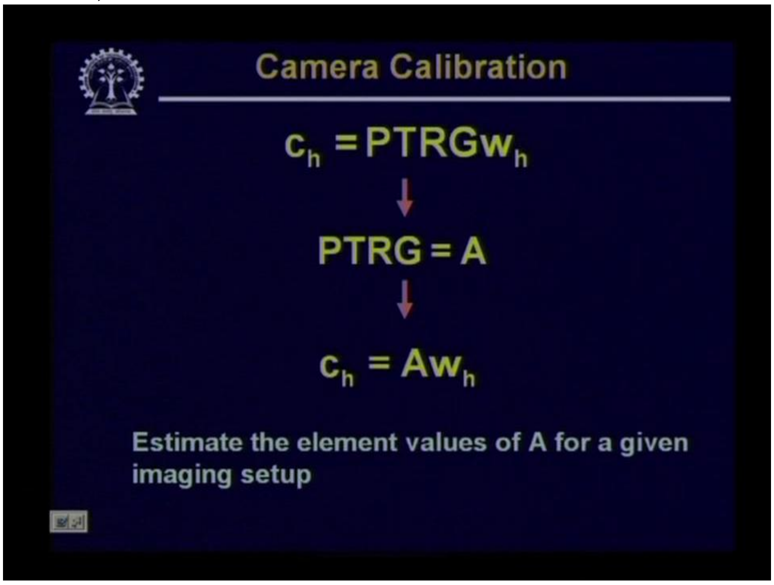
So it is possible that for a given camera for which

(Refer Slide Time 11:31)

the focal length lambda is known I can find out what is the corresponding perspective transformation matrix p where as to find out the other transformation matrices like t, r and g I have to do the measurement physically that what is the translation of the gimbal center from the origin of the 3D world coordinate system, what is the pan angle, what is the tilt angle. I also have to measure physically that what is the displacement of the image center, image plane center from the gimbal center. And in many cases measuring these quantities is not very easy and it is more difficult if the imaging setup is changed quite frequently. So in such cases it is always better that you first have an imaging setup and then try to calibrate the imaging setup with the help of the images of some known points of 3D objects that will be obtained with the help of the same imaging setup.

So by calibration, what I mean is as we said that now I have a combined transformation matrix for the given imaging setup which is which is nothing

(Refer Slide Time 12:48)

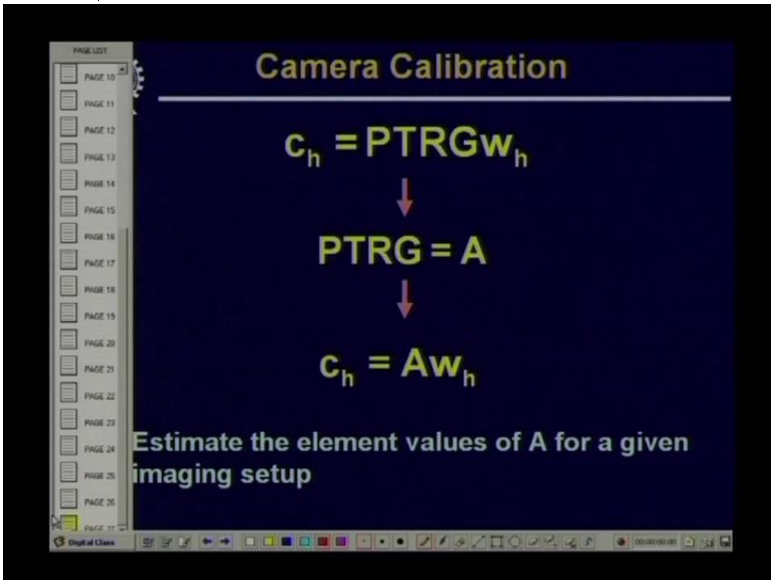


but the product of p, t r and g So this being a 4 by 4 matrix what I have to do is I have to estimate the different

(Refer Slide Time 12:59)

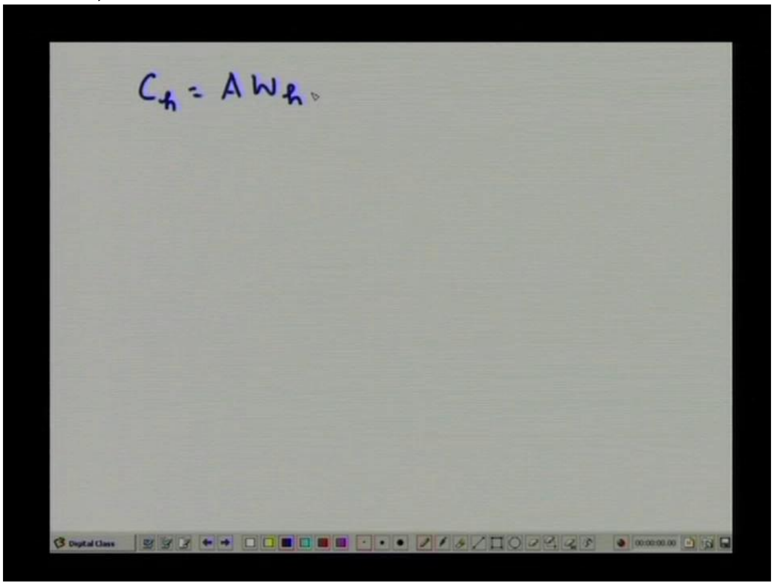
element values of this matrix a So if I can estimate the different element values of the total transformation matrix a from some known images, then given any other point in the 3D, I can find out what will be the corresponding image point. Not only that if I have an image point, a point in the image, by applying the inverse transformation I can find out what will be the equation of the straight line on which the corresponding world point will be lying. So this calibration means we have to estimate the different values of this matrix a. Now let us see how we can estimate

(Refer Slide Time 13:43)



these values of the matrix a So here you find that we have this matrix equation which is of this form that is c h is equal to a into w h where we have said

(Refer Slide Time 14:07)

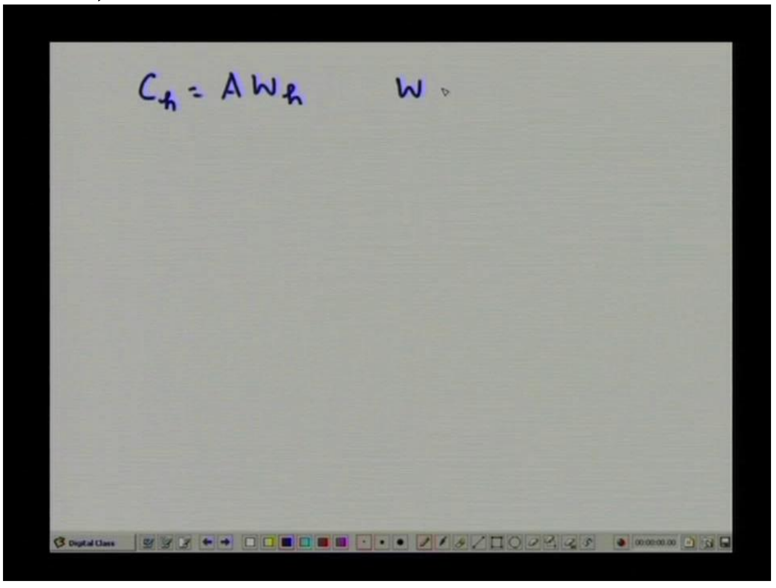


that w h is the world coordinate of the 3D point

(Refer Slide Time 14:12)

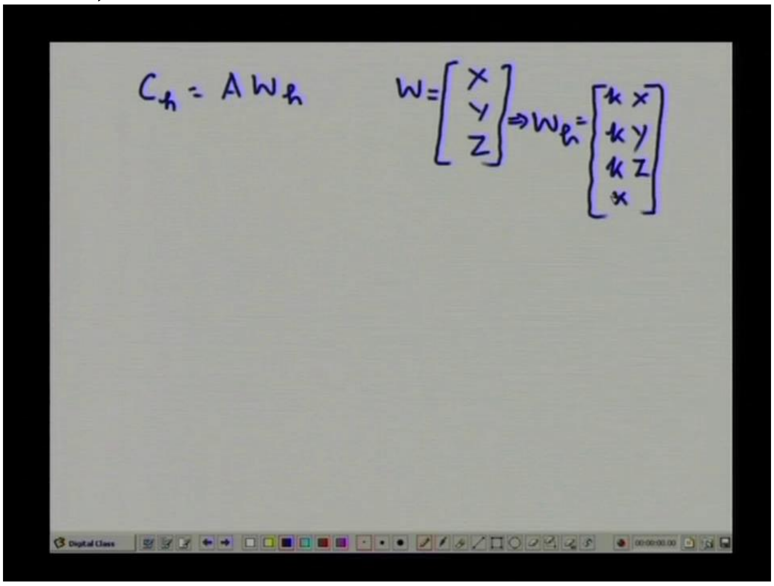
put in homogenous form and c h is the image point on the plane again in the homogenous form and a is the total transformation matrix So here if the world point w

(Refer Slide Time 14:27)



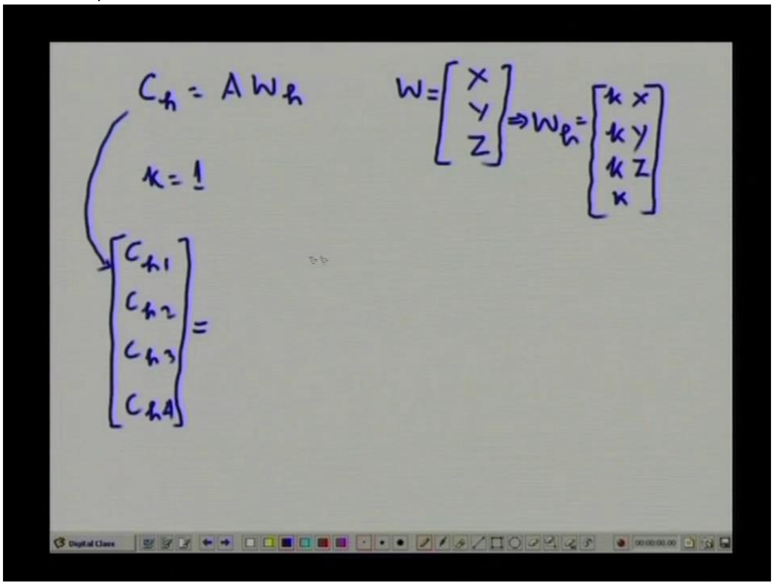
has the coordinates say x y an z, the corresponding homogenous coordinate system will be given by w h is equal to some constant k times x, some constant k times y, some constant k times z and the fourth element will be k. So this will be the

(Refer Slide Time 14:53)



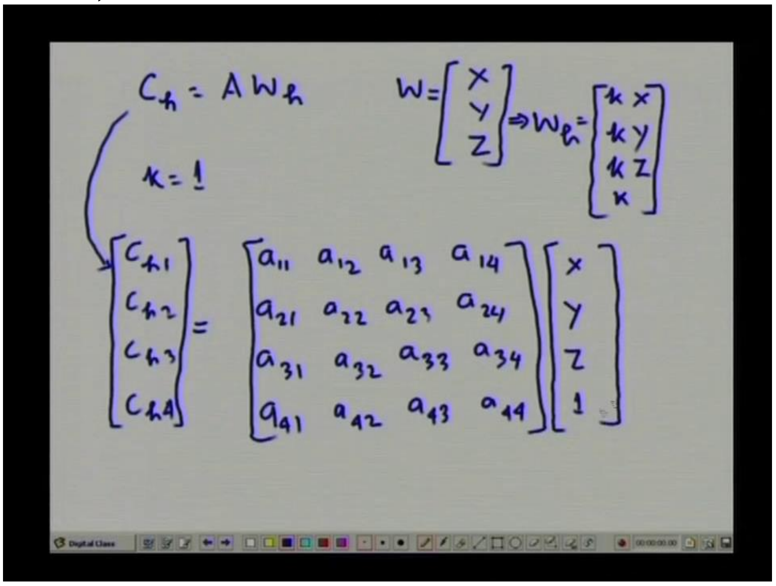
homogenous coordinate w h, corresponding to the point w Now without any loss of generality I can assume the value of k equal to 1 So if I take k equal to 1 and if I expand this matrix equation then what I get is I get the component c h 1, c h 2, c h 3, c h 4 this will be, now I expand the matrix a also

(Refer Slide Time 15:28)



so a will have the components a 1 1, a 1 2, a 1 3, a 1 4; a 2 1, a 2 2, a 2 3, a 2 4; a 3 1, a 3 2, a 3 3, a 3 4, then a 4 1, a 4 2, a 4 3, a 4 4 into the homogenous coordinate of the point in the 3D space which is now x y z and 1. So you remember that we have now assumed

(Refer Slide Time 16:21)



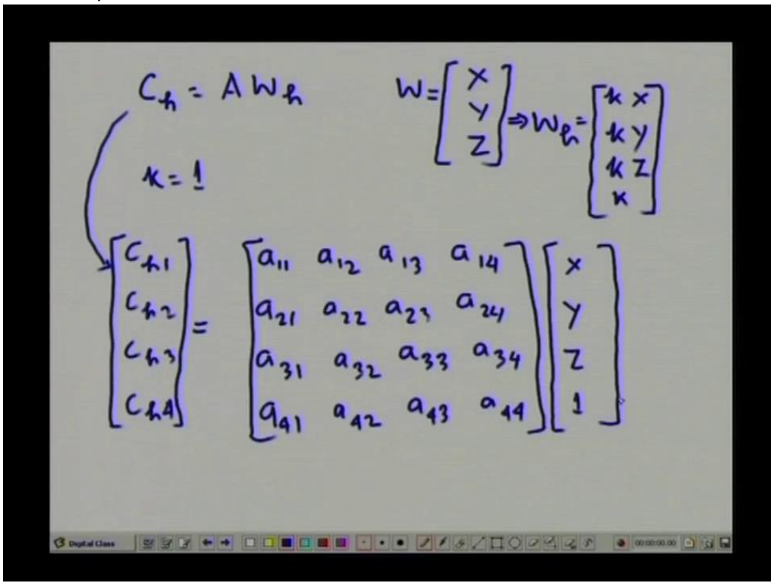
the value of k to be equal to 1 So I get a matrix equation like this

Now from this matrix equation I have to find out

(Refer Slide Time 16:31)

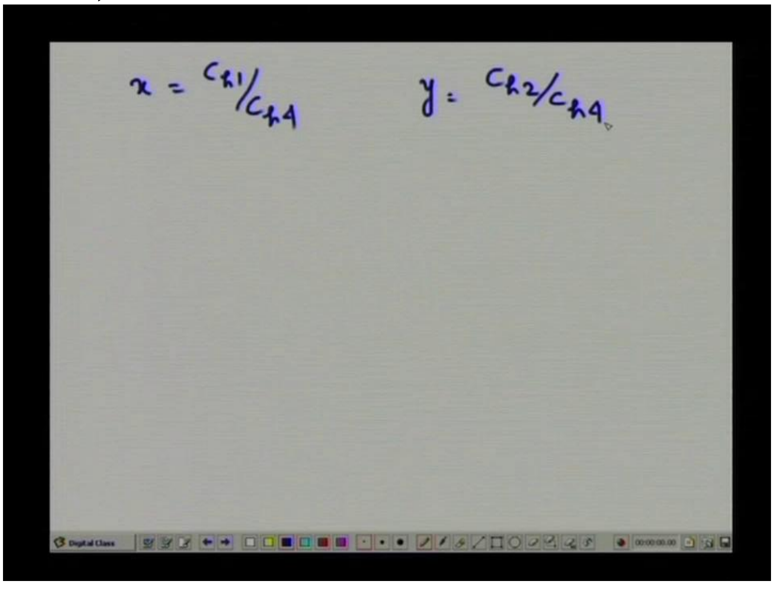
or I have to estimate the component values a 1 1, a 1 2, a 1 3 and so on. Now here

(Refer Slide Time 16:39)



once I have the homogenous image coordinates c h 1, c h 2, c h 3 and c h 4, then we had already discussed that the corresponding Cartesian coordinate in the image plane is given by x equal to c h 1 divided by c h 4 and y is given by c h 2 divided by c h 4. So this is simply a conversion

(Refer Slide Time 17:09)

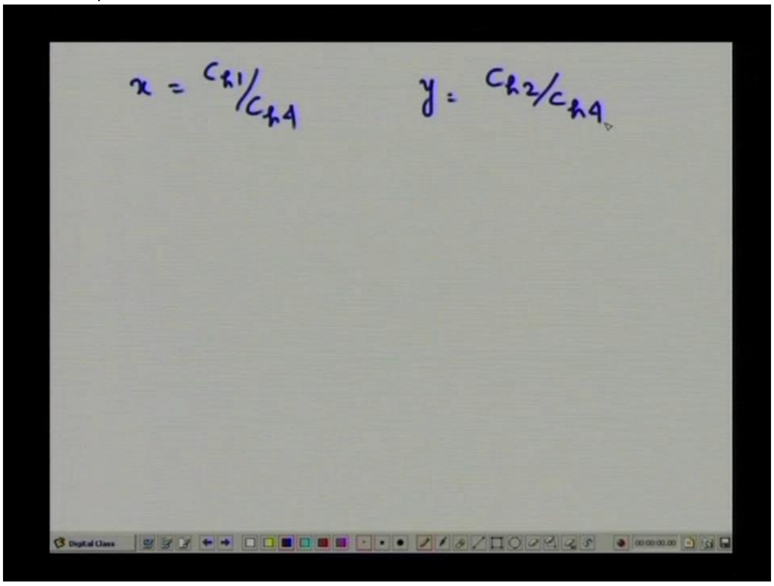


from the homogenous

(Refer Slide Time 17:11)

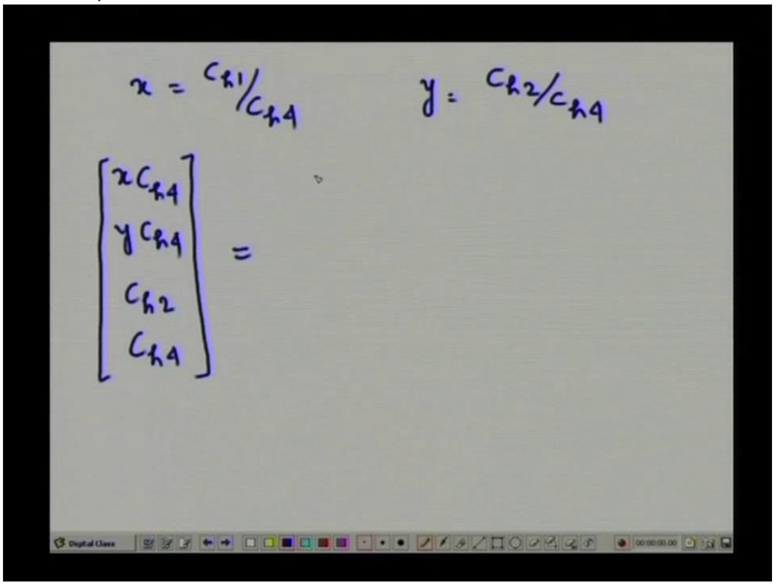
coordinate system to the Cartesian coordinate system. Now here, if I replace the values of c h 1 and c h 2

(Refer Slide Time 17:23)



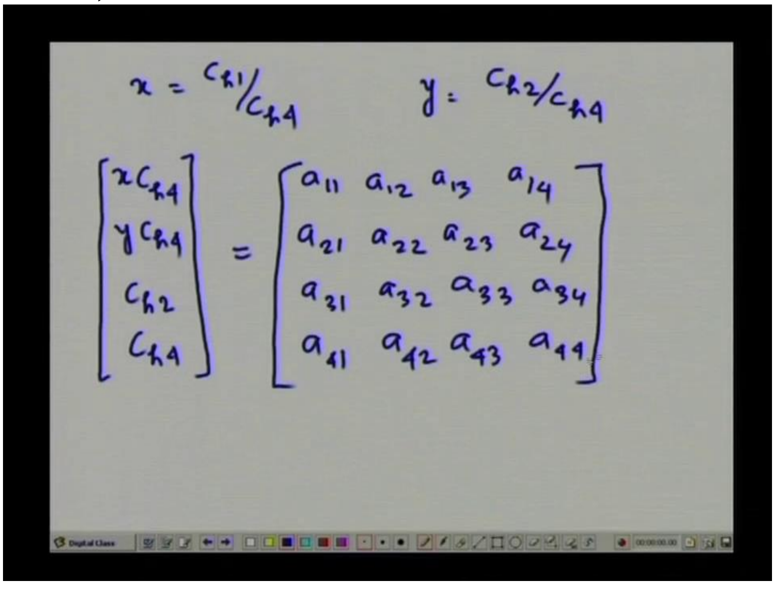
by x times c h 4 and y times c h 4 in our matrix equation then the matrix equation will look like x c h 4, y c h 4, then c h 2 let it remain as it is and finally we have c h 4 this will be equal to

(Refer Slide Time 17:53)



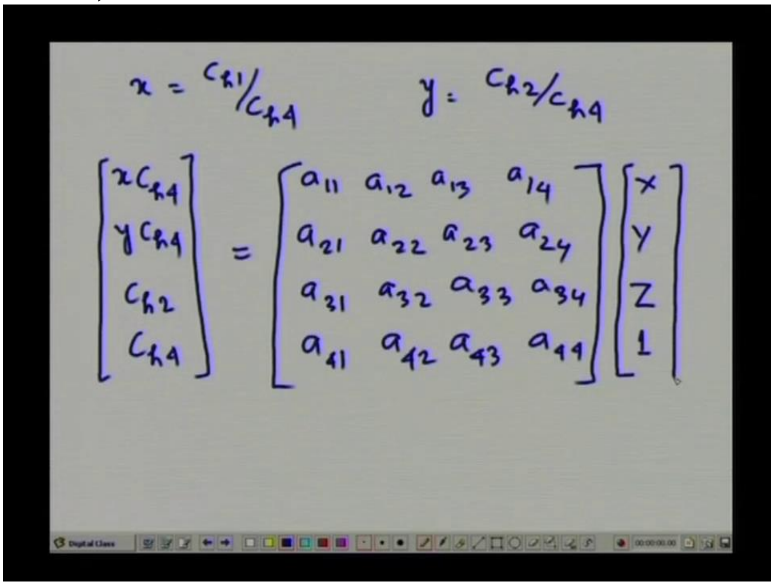
a 1 1, a 1 2, a 1 3, a 1 4, a 2 1, a 2 2, a 2 3, a 2 4, a 3 1, a 3 2, a 3 3, a 3 4, a 4 1, a 4 2, a 4 3, a 4 4 multiplied by the 3D

(Refer Slide Time 18:26)



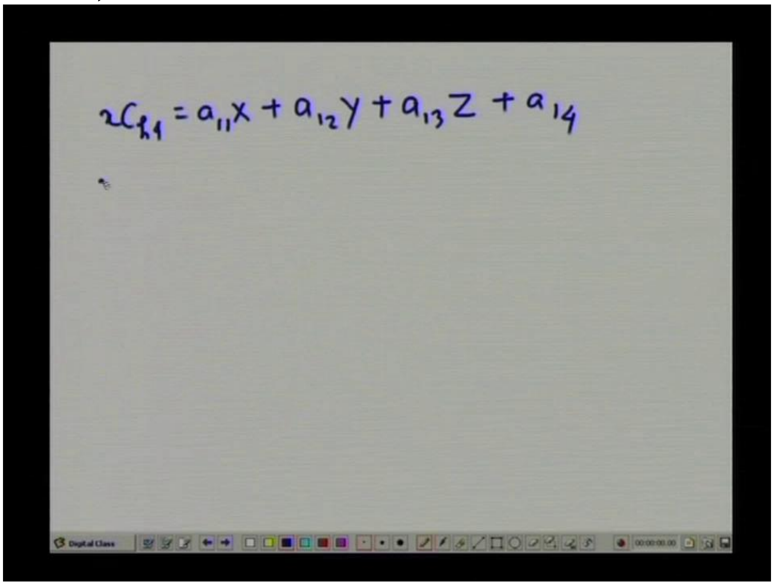
point coordinate in homogenous form which is x y z 1

(Refer Slide Time 18:36)



So if expand this matrix equation what I get is x c h 4 will be given by a 1 1 x plus a 1 2 y plus a 1 3 z plus a 1 4,

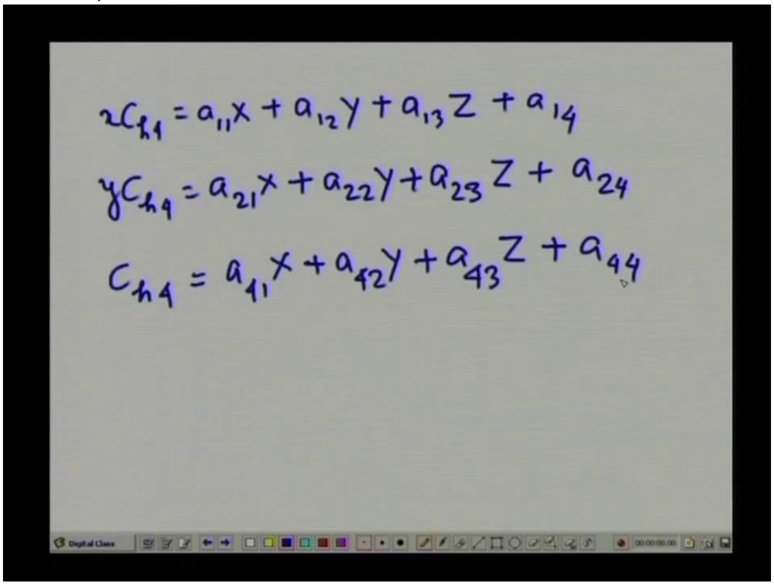
(Refer Slide Time 19:03)



then y c h 4 will be equal to a 2 1 x plus a 2 2 y plus a 2 3 z plus a 2 4 and c h 4 is given by a 4 1 x plus a 4 2 y plus a 4 3 z plus a 4 4.

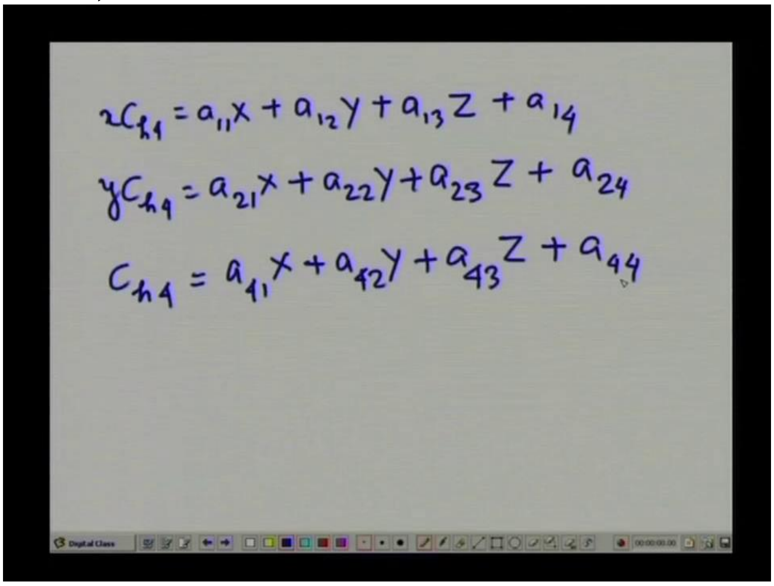
Now find

(Refer Slide Time 19:40)



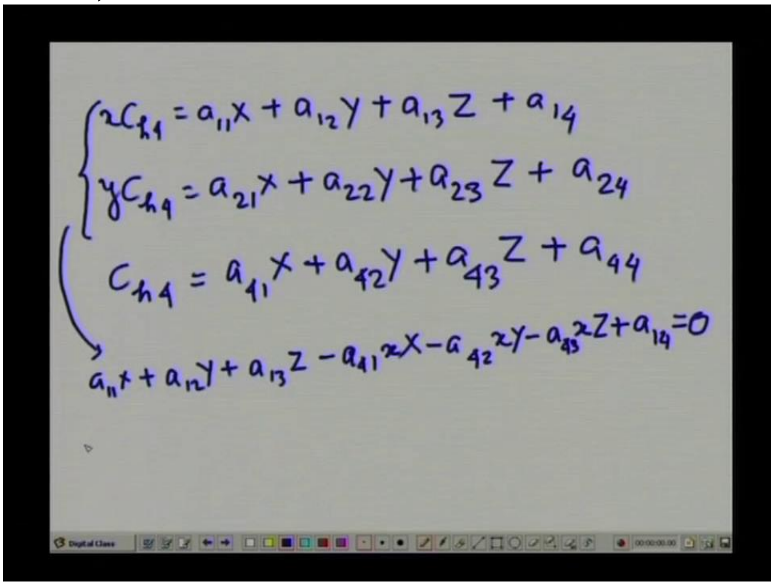
that while doing this matrix equation or while trying to solve these matrix equations, we had ignored the third component in the image, image point That is because the third component corresponds to the z value and we have said for this kind of calculation the z value is not important to us. Now from these given 3 equations, what we can do is

(Refer Slide Time 20:10)



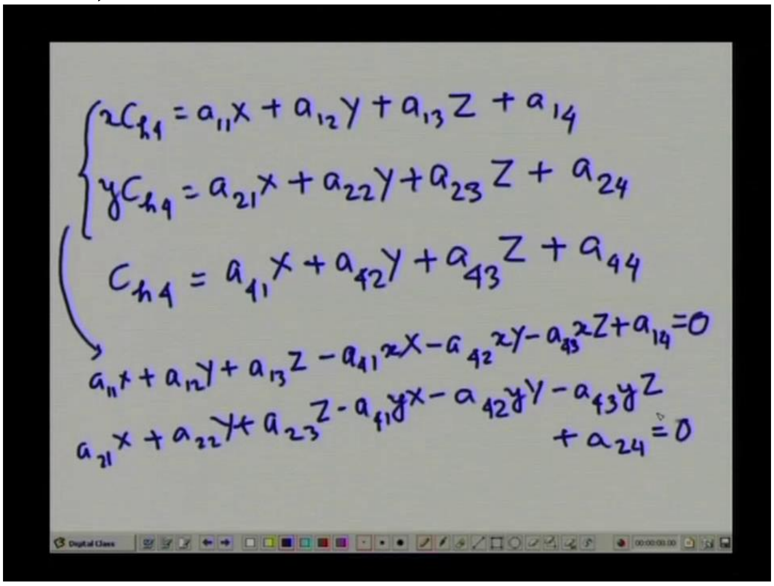
we can find out what is the value of c h 4 in terms of x y z and if replace these values of c h 4 in the earlier two equations then these two equations will simply be converted in the form a 1 1 x plus a 1 2 y plus a 1 3 z minus a 4 1 x small x capital x minus a 4 2 small x capital Y minus a 4 3 small x capital Z plus a 1 4, this is equal to 0 and

(Refer Slide Time 21:02)



a 2 1 capital X plus a 2 2 capital Y plus a 2 3 capital Z minus a 4 1 small x, small y capital X minus a 4 2 small y capital Y minus a 4 3 small y capital Z plus a 2 4 this is equal to 0.

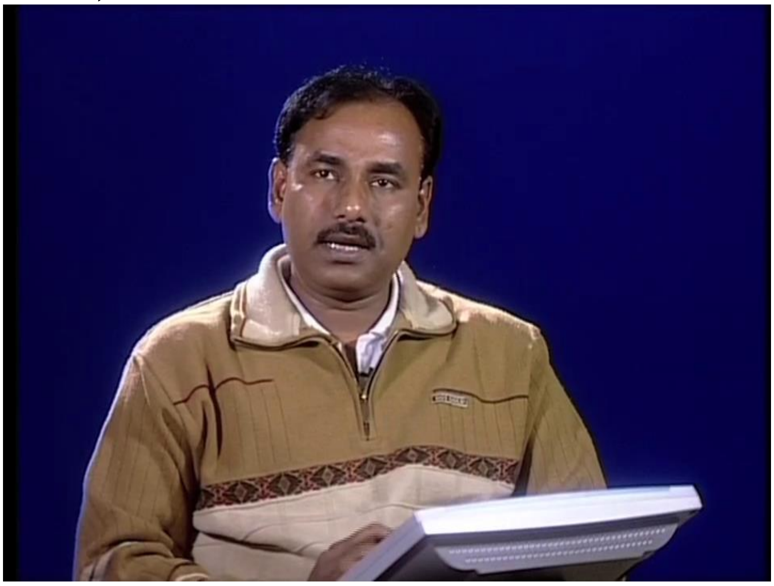
(Refer Slide Time 21:43)



These two equations are now converted in this particular form.

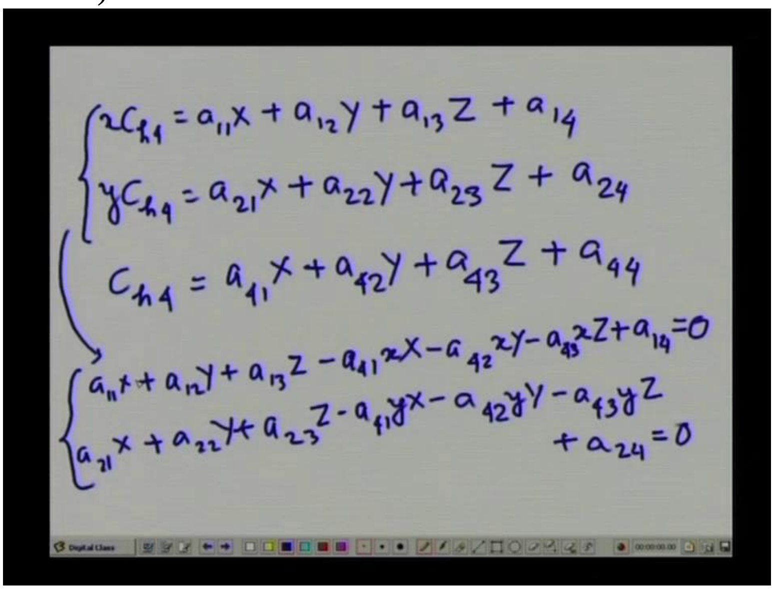
Now if you study these two equations you find that

(Refer Slide Time 21:53)



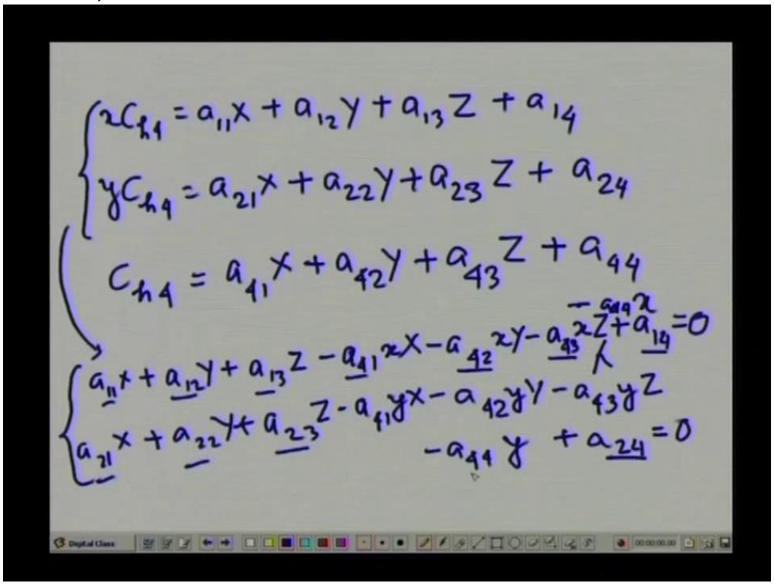
x and y, small x and small y are the coordinates in the image plane of a point in the 3D world coordinate system whose coordinates are given by capital X capital Y and capital Z. So if I take a set of images for which a point in the 3D world coordinate system, that is capital X capital Y and capital Z are known and also find out what is the corresponding image point, image coordinate in the image plane, then for every such pair of readings I get 2 equations, one is the first equation, other one is the second equation. Now if you study this particular, these two equations

(Refer Slide Time 22:37)



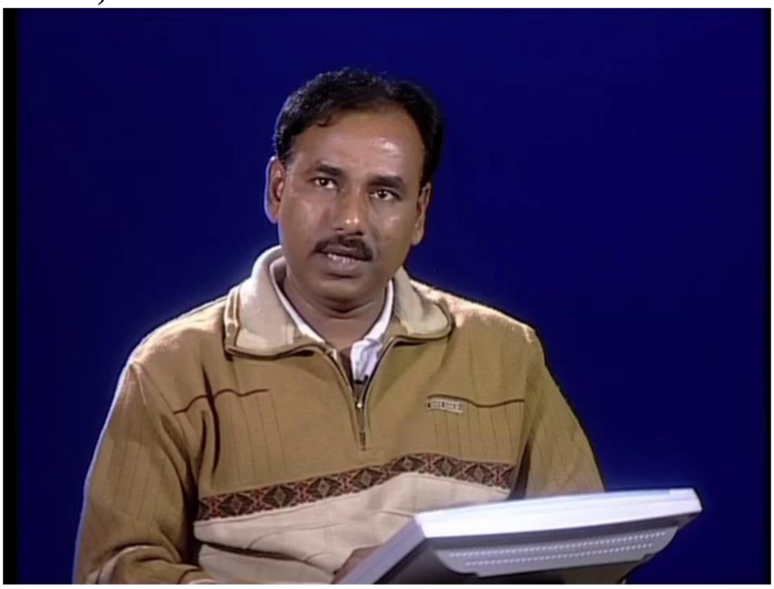
you find that there are 6 unknowns. The unknowns are there is a 1 1, a 1 2, a 1 3, a 4 1, a 4 2, a 4 3, a 1 4, a 2 1, a 2 2, a 2 3 then you have a 2 4. So the number of unknowns we have in this equation are one, two three, four, five, six, seven, eight, nine, ten and eleven. So there, eleven or twelve; one, two three, four, five, six, seven, eight, nine, ten have I missed something sorry there should be one more term minus here there should be one more term minus a 4 4 x and here should be one more term minus a 4 4 y, so this a 4 4 this is another term

(Refer Slide Time 23:59)



so there are twelve unknowns. So for solving these 12 unknowns, we need 12 different equations and for every known point in the

(Refer Slide Time 24:08)



3D world, I get two equations. So if I take such images for 6 known points then I can find out thank you.