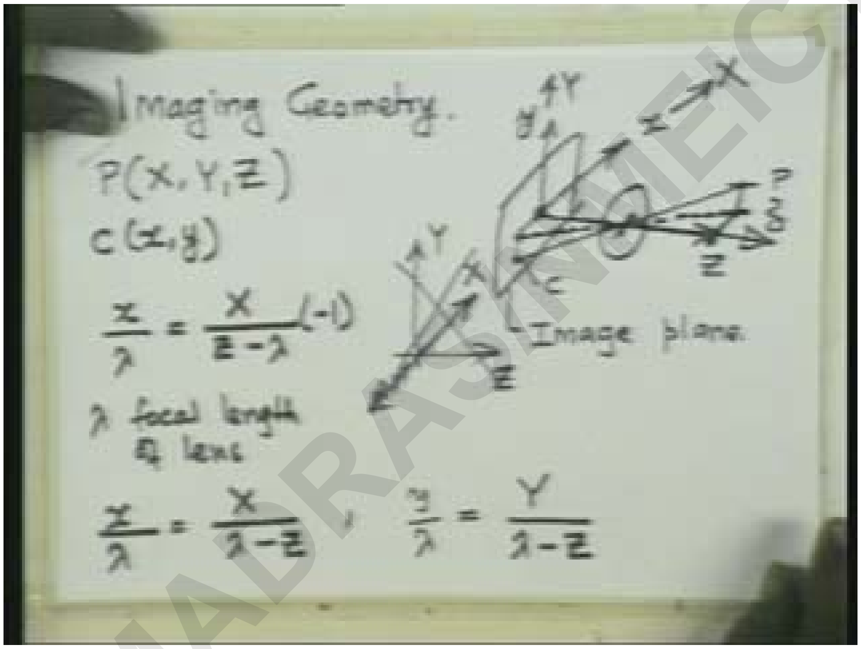
ROBOTICS

Prof. B.Seth
Department of Mechanical Engineering
IIT Bombay

Lecture No-26

Image Processing

We look at some basic aspects of robot vision in particularly we try to see what the major tasks are in robot being able to use information from a camera type of imaging system today we will start out by looking a the imaging geometry or we deal with transformation that is there between [noise] and the camera coordinates so basically(refer slide time07:57)



what we have to look at it is that we have a lens of the camera and behind that there is a plane on which image is focused

so if i draw the image plane so that form the center of this to the center of the lens is what i will define the axis of the lens or the camera axis and lets call that is called as small z now let us define the coordinates to have x and y in the plane of the in the image plane right so this is the image plane

if i take it coordinate a point on the whole something outside of the lens obviously so i am looking at the coordinate with some global oops lets keep the axis the same so this is my capital z axis this is my capital z axis this is my capital x axis and this is my capital y axis ignore this i am taking infact what i like to do is to take this coincidental with that right so i will take my capital z axis capital x axis and capital y axis coincidental with the coordinates which i define with respect to the image plane okay

so this also you can ignore now if i if this point p has coordinates which are given in the global frame of reference as x y and z what we are interested in knowing is what this point corresponds to in the image okay now if we assume optics to be simple then basically if you take a point if this has got some co ordinates x y and z so this is the z coordinate from the beginning here to here and then we have x coordinate which is the distance from here to here and y coordinate which is the distance from here to here along the vertical direction this is along the x direction so this is the coordinates x y and z this is the lens center and i will get a image point which is going to have some x which is going to be inverted and some y which is going to be inverted give me a point here which let me call this point as c

okay this is the camera image point this is the world point now what i want to know is what is the relationship between the and the coordinates of this point and the world coordinates world coordinates are x y and z and the camera coordinates are small x and small y right

so relationship is very clear if you look at if i take this particular point here and join it then i will get two triangular two triangles which are going to be similar triangles right because this angle is the same as that angle and then the triangles are similar and therefore i will be able to write relationship which is small x divided by the distance from the focal plane or the image plane to the center of the lens

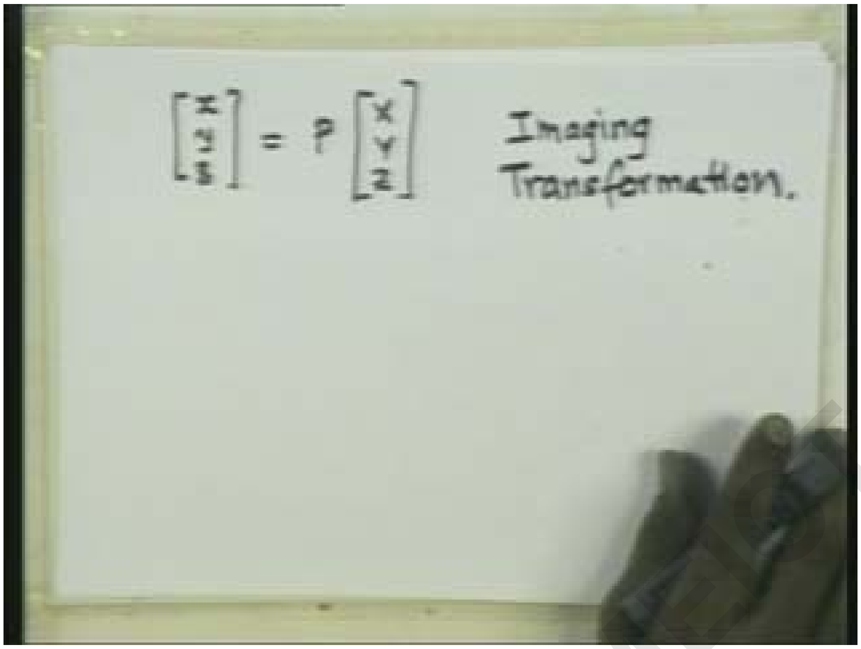
this distance is called the focal length of the lens okay so that let us denote focal length as lambda so lambda is focal length of lens okay

so this is now equal to what if this is one triangle here so this is the triangle that i am looking at and so i have to look at the x coordinate here okay

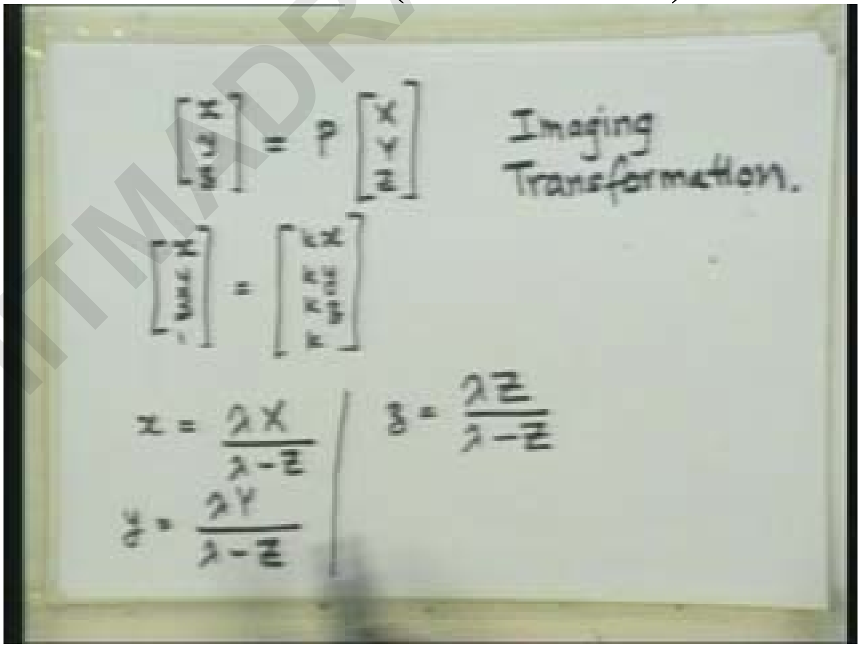
so that is equal to capital x divided by this distance from the center of the lens to the z coordinate right

so that is going to be z minus lambda but you please notice that when x is positive here since these two coordinates are coincidental i will get negative x here right so i have to take this with a minus sign in other words i should i should write this as x over lambda equals to capital x divided by lambda minus z

is that clear its basically simple you have similar triangles which i am dealing with and i have written down ratio of two sides of similar triangles now similarly i can look at the y coordinate here by looking at this triangle and corresponding triangle here which are again similar triangles and i should be able to write y over lambda equals to capital y over lambda minus z right i got this now if i want to proceed further i would like to now try to define x y z where z is zero [noise](refer slide time08:46)



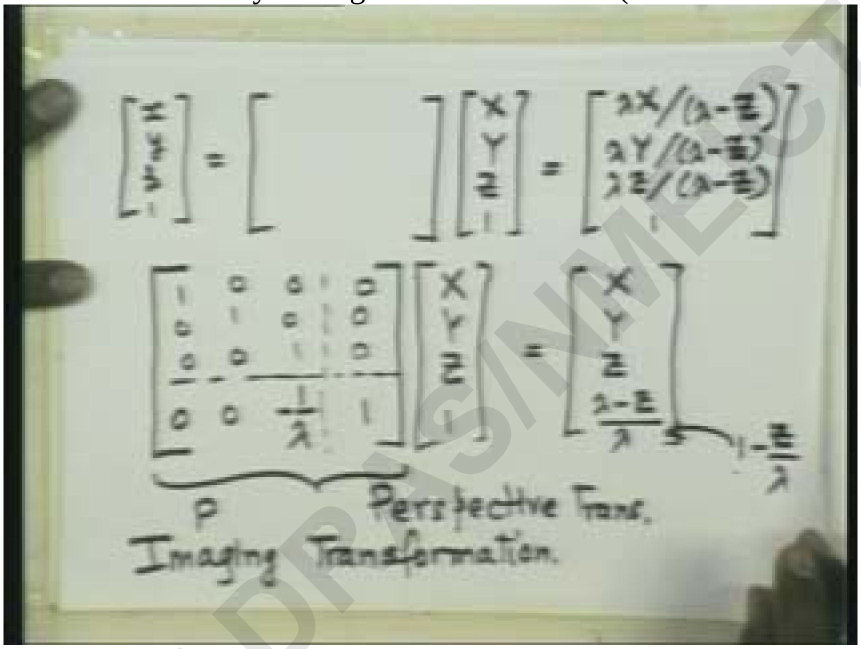
as some kind of transformation times may global coordinates x y and z
i would like to have some kind of a imaging transformation okay now you have already learned about homogeneous coordinates right in dealing with the coordinates of motion of robot frames so actually we will use the homogeneous coordinates because they will turn out to be useful here also and you just have to remember that if i have regular coordinates like this i can define homogeneous coordinates which are going to be x y z and unity which is taken to be the same as(refer slide time10:34)



k times x k times y k times z and k right these two coordinates are considered to be the same in other words anytime in the homogeneous coordinate system if you want to get the real coordinates you divide it by the last element of the vector okay so now what we have so far is we have x is equal to lambda times capital x divided lambda minus z this is the relationship i had and i had this relationship right what i am going to do is i will do a little trick and i will say what about this right i have written a relationship which is similar to what i have written for x and y for the z coordinate

now we know that z ordinate is actually zero because we are talking about the image plane okay however it is jelling with whatever i have written for x and y coordinate so it turns out that it is useful to do that although it is not meaningful in terms of interpretation but it gives me a pre variable which will allow me to define a transformation which will work for the imaging geometry that means given a world point how do i find out what is the image point that is the key question thatI am trying to address

okay then let me write down my homogeneous coordinates (refer slide time15:15)

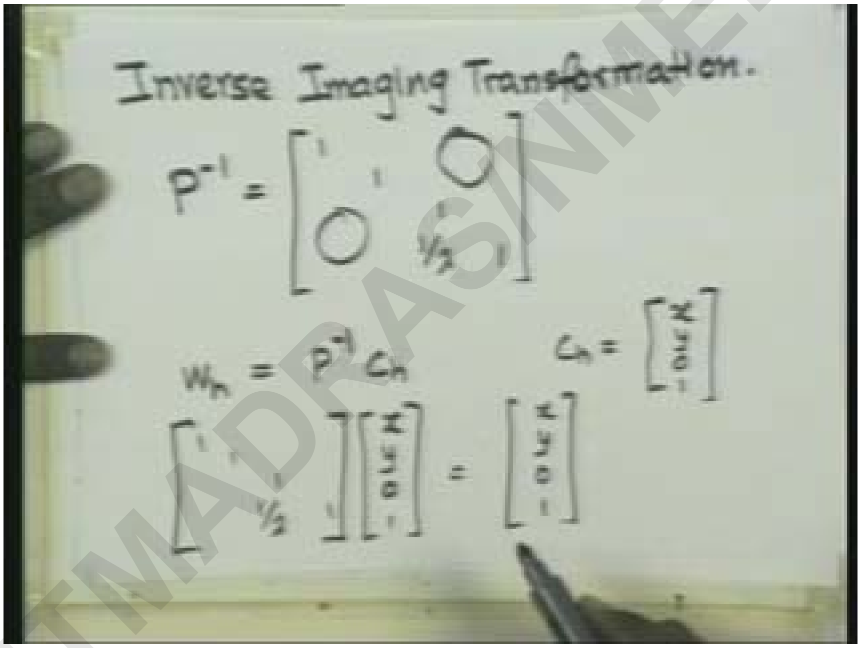


i am writing as x y z and one and what is going here is what i am interested in knowing what i know already is that if i apply to the global coordinates x y z and one i should get this quantity but i also have found out that this is also equal to lambda times x divided by lambda minus z right here i have lambda times y divided by lambda minus z here i have lambda capital z divided by lambda minus z and i have one here

okay so now what this i can see that i can divide all the coordinates here by the same because these are homogeneous coordinates so what i like to do is i will like to put only x y and z here and take out whatever is common as the last element so now i have on the rightern side this vector here can be written as x capital that is capital x, capital y, capital z and i have lambda minus z divided lambda

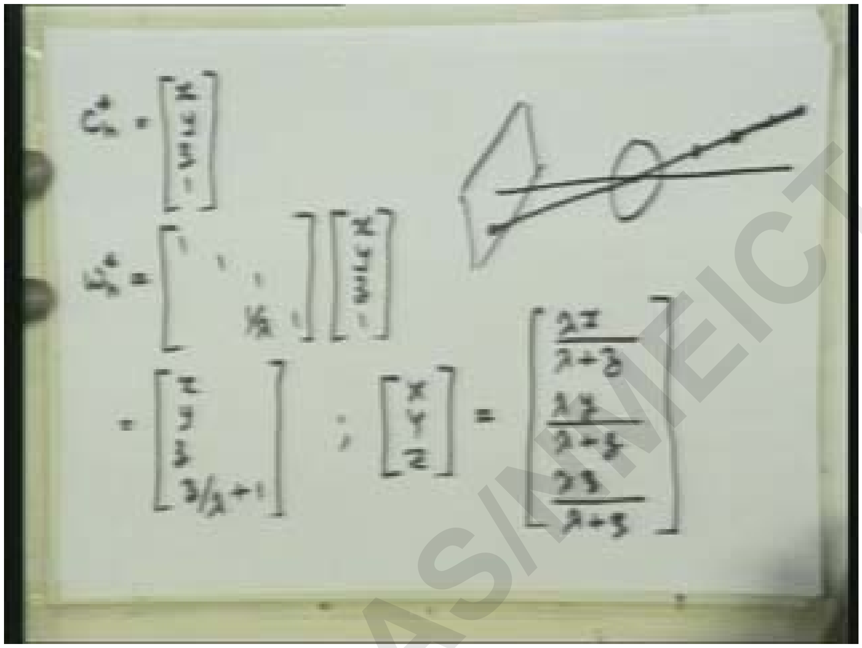
you agree with me on that basically this is homogeneous coordinates so whenever we can take a common factor out from here and put it in the as the last element okay the last element here is basically if i look at this element its nothing but one minus z over lambda okay now let me once again put my unknown p this is my global coordinates of the point of interest right away i see that first three elements in these two vectors are the same

so what goes in this matrix is clear right It has to be nothing but the identity matrix so i have a sub matrix here which is got identity matrix so you have all only diagonal one elements right
so i have already with these ones i will get these on the rightern side so these elements are also zero right only thing now remaining is i have on the right rightern side last element is one minus z over lambda i have got capital z here i have got one here
so if i have a one here then i will get this one i need to get minus z over lambda here i have z here so if i put one over lambda with a minus sign as this element and put zero zero here you can convince yourself that actually you are going to get in the rightern side and this here i will call as the imaging transformation okay
so also called perspective transformation it is very useful for imaging and actually finding out how the perspective of a point world coordinate looks like okay
so we have now defined the imaging transformation the question is if i have to find it useful can i do tricks with it like can i find the inverse of the imaging transformation okay so i am(refer slide time17:50)

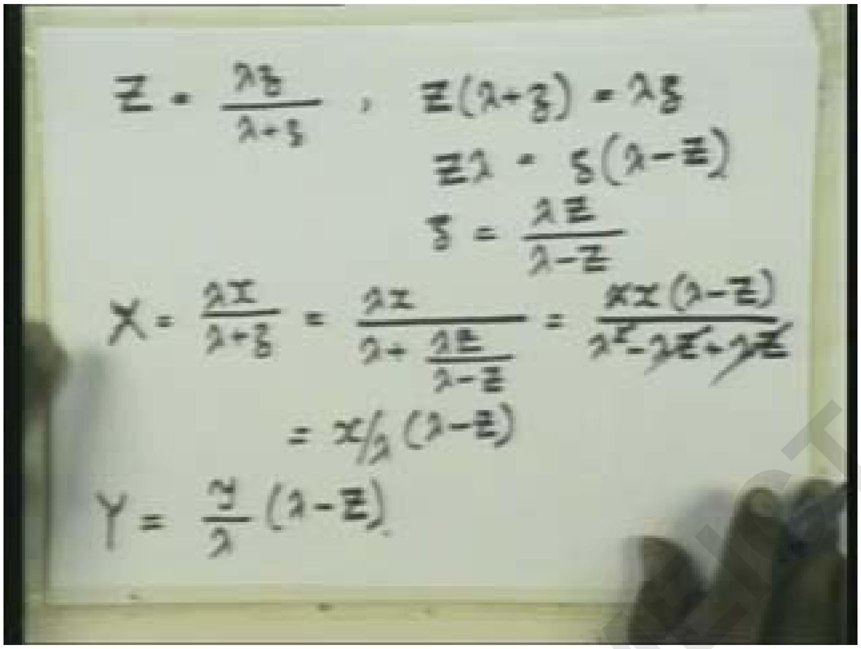


interested in knowing what p inverse is given that what we have for p you can show it that p inverse is nothing but i am only showing you the non zero elements of the four by four matrix all other elements are zero you can multiply out p with p inverse and you will see that you will get identity matrix okay
so that is very nice because if i have a p inverse what it relates is that if want to get the world coordinates homogeneous coordinates of the world some point all i have to do is to take inverse imaging transformation and multiply it with the camera homogeneous coordinates right
so if we do that let us apply that let us now say that camera coordinates are some x y zero one and i have my p inverse so i have p inverse just got one one one one and one over lambda and i am multiplying this x y zero one what i will get here is basically x y zero and the last row is going to be one also okay

so what it is told me is that the world coordinates and the camera coordinates are coincidental which is not very useful right the problem of course is very clear what we are trying to deal with is as we have we know already that image is nothing but the reflection of the three dimensional world on to the two dimensional image plane
so obviously you are losing some information you are losing information so you cannot automatically say what is the point so you know you have the image plane(refer slide time20:40)



which has only limited amount of information if i take some point on the image plane i know that the world coordin point must be lined here here here here or anywhere ofcourse mathematically it could lie here also and that is what it is telling me so what is telling me is not wrong but it is not very useful
so what we need to do is once again do a little bit of trick and say okay let me assume there is some z coordinate for the image point also i dont take it as zero i will take it as z so let us now apply my so i am saying camera some modified version of that is having x y z and one now If i apply to find out my world coordinate then i am using p inverse matrix which is one over lambda one times the full set up coordinate x y z and one and what i will now get is x y z and the last row is z divided by lambda plus one and from this if i wanted to get the x y z locations of the world coordinates
so If i want to get x y z of the point of interest whose image i have then i can divide it by the last row to get me lambda x over lambda plus small z lambda y over lambda plus small z lambda z over lambda plus small z okay
now this has more potential for giving us little more useful information what i will do is i will take the last equation these are three equations essentially
so If i look at the last equation i will have a relationship which can be useful okay so the last equation that i have (refer slide time22:56)



is capital z equals to small lambda lambda and small z divided by lambda plus small z if i manipulate that this is the same as capital z times lambda plus small z okay equals to lambda small z i will collect all terms of small z together i will get z times lambda equals to small z times lambda minus capital z

therefore small z is nothing but lambda z over lambda minus z okay this is this is what we had got earlier also okay there is no difference now i have expression for capital x as small x times lambda divided by lambda plus small z and i substitute for small z from what i have got so this is equal to lambda times x divided by lambda plus small z which is lambda z divided by lambda minus z

okay and this can be written as multiplied and divided by lambda minus z so i get lambda x times lambda minus capital z divided by lambda squared minus lambda capital z plus lambda capital z these two terms cancel out and i can divide both sides by lambda give me x divided by lambda times lambda minus z right this is simple algebraic manipulation i have done similarly without actually carrying it out i can confidentally write that this is equal to that

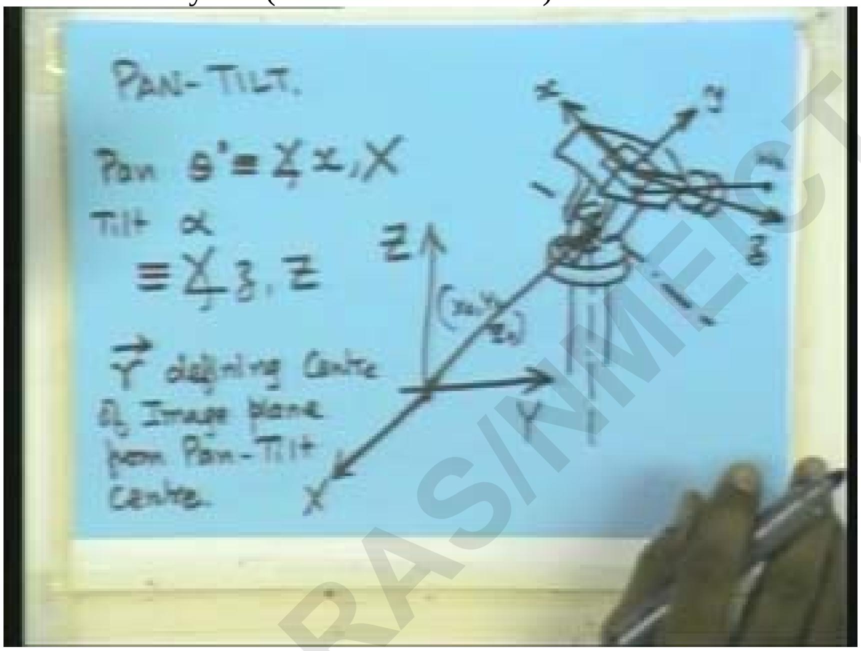
okay what i have got here is what is intuitive what i had shown here is that given here a camera coordinate there can be any number of world coordinates so what we need is we need some more information the information that we need is given here if i know the coordinates the z coordinates of the world point then i am able to actually determine the x and y coordinates of the world point right the additional piece of information i need it does not have to be z it could have been x it could have been y and then we could have done similar thing and found out the unknown small z and then we could have substitute it to get the same thing right

so that additional piece of information is necessary okay so now we are going to look at situation where we may not have the camera in the same spot as the imaging plane or world coordinate system may not coincide with the camera coordinate system

you have a question um how do you determine small z we just leave it as a pre variable there is a relationship that is there between the small z and the capital z so eventually it boils down to the fact that you need to know the capital z for determining capital x and capital y thats what it turns out

no because there is not a there is a imaging transformation is a many to one transformation that means many points in the real world are going to coincide with the same point on the image its a two dimensional image three dimensional world right

so anything along a single line is going to map to the same point so when you want to find the inverse map there is no way you can actually determine the other point that is what we have okay so the situation that i am now talking about let me put that down we have a global coordinate system(refer slide time31:32)



okay now imagine that we have a camera which is somewhere else here is the lens right so the camera axis is and i have to show image plane somewhere inside the camera i have the image plane and that image plane has a coordinate system attached to it so this is my small x small y z i am always taking along the axis of the lens

now what i am imagining is that many times here situation where this is going to be mounted on a what is called a pan tilt arrangement so i have a pivot here

okay you get the idea i have a axis of rotation which is vertical with respect to the global coordinate plane and i have a axis which is i am going to assume this to be parallel to the small x axis so it is along the image plane so what you do is typically if you rememb if you recall you know seeing studios people will have cameras and they will have or we have cameras right here there will be a handle you release the handle you can pan it you can you know scan the whole of the view horizontally or you may be able to tilt it by lifting it up and down the handle

so that you can look down or up right this is called pan tilt arrangement so with this pan tilt arrangement now what we are interested in is because in many robotics applications also the camera will be at a different spot the world coordinate system will be somewhere else infact there will be situations where you you will mount the camera on the robot itself so not only you have the two different but the camera will be actually moving with respect to time and it will be occupying different spots in the global frame

so if you are able to deal with this type of situation we will feel confident dealing with the situation where the camera would be mounted on the you know robot or will be else where what we are interested in is given a point in the real world this is our world homogeneous coordinate which will correspond to if i take this is the center of the lens this is getting flattered up somewhere here i am going to have a image point on the image plane i want to know what the x and y coordinates of that is for the point which is given in the global frame of reference what obviously i need to know is where the camera is located how much it is panned how much it is tilted right

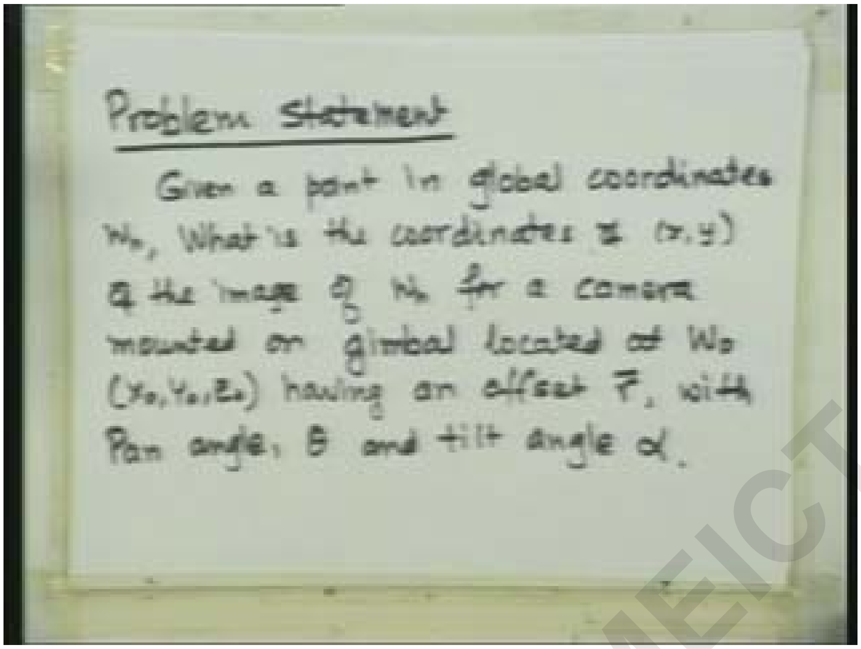
so we need to know this vector here what i am going to do is i am going to look at the centre of pan tilt is got intersecting axis this is the vertical axis there is a x axis along the x axis wherever it intersects i take that point and let me take a vector from here to global coordinates to that that i will define as one vector which defines where the camera is then i will i am going to say there is some pan and tilt pan is what is about this vertical axis so pan angle is theta so let me first say that this is the camera location

so it has coordinates x not y not z not this particular point of intersecting axis then i am saying that let us have some pan of theta degrees angle theta and tilt of alpha how do i define pan and tilt well pan is what pan i am rotating about the vertical axis which is coinciding with the global z axis

so if i say the angle between the capital x the global x axis and the camera image plane x axis so i will define this as the angle between x and capital x similarly i will define this as the angle between capital z and the small z because this is the rotation tilt is the rotation about this axis which is parallel to the small x axis right

so this is going to be the angle between the small z and capital z if i define it like that then it is uniquely going to find out and then i have to know one other vector which is from this center of the pan tilt to the center of the image plane so that let me define it as vector r okay this is defining centre of image plane from pan tilt centre

is the geometry clear to you okay so now what do we problem that we are trying to address is (refer slide time34:42)



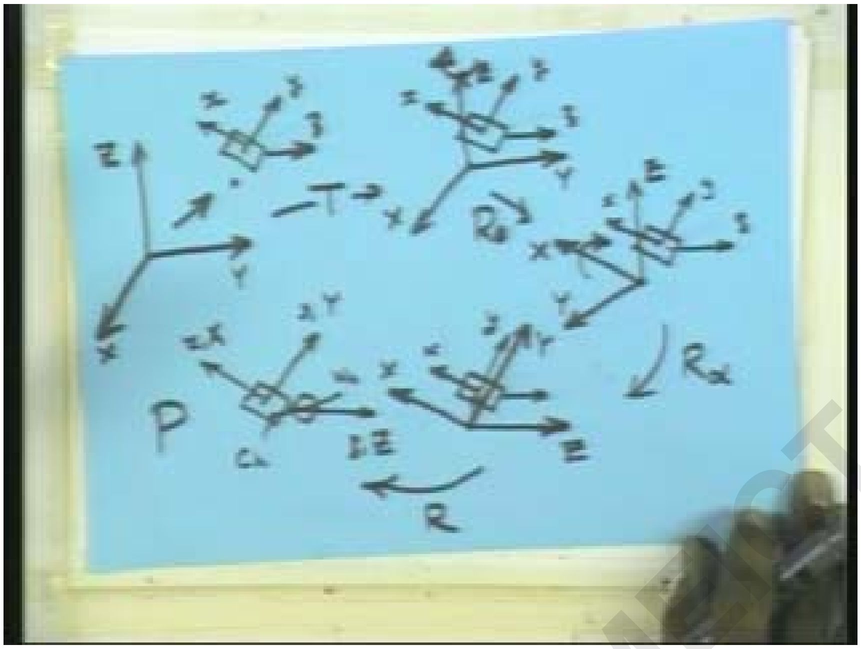
okay given a point in space a point in global coordinates okay say w h what is the coordinates

okay x y small x y of the image of point w h okay for a camera mounted on this called gimbal mount okay gimbal located at w zero okay that has the coordinates x zero y zero z zero okay having an offset r okay with pan angle theta okay as i have to find already and tilt angle alpha okay this is my problem that i am trying to address i have already when i talked about the imaging transformation i had assumed that the coordinates of the global frame and the camera frame were the same okay

so now i am saying that i dont have to restrict myself to that and i am going to let them vary i am i am going to have different coordinates for the camera and different coordinates for the global frame

so how do i now find that imaging transformation that i am looking for okay the way to go about this is to say we already know how the transformations are going to be if i allow my global frame to be translated or rotated until it becomes coincident with the camera coordinates if i can do that then i have already got a perspective transformation or imaging transformation and then i can put the imaging transformation and if necessary throw it back

so that is what i wanted to do now for doing that let me try to illustrate this because this is going to be important to understand how this transformation is obtained so let us let me attempt this okay i will have to try a series of figures so this is my global coordinates(refer slide time41:21)



x y z i am not going to show the whole camera but i will show you the image plane the center of the image plane here axis which is the small x small z and small y and small x okay this is what i have shown you in the big picture now i am showing this in a skeletal form the two coordinates are shown this is the image plane lens of course will be lens center would be on the z axis small z axis
so first thing i want to do is translate this global frame so that it coincides with the pan tilt point pan tilt point is the intersection of the pan axis and the tilt axis right
so if i can translate this center here then i will be able to bring my coordinates so if i draw that now after translation what i will have is this point this image plane is in the same place it has its z axis y axis and x axis okay i am let me write this if this is not going to flutter it and now i have shifted this big global frame to this point here
so this is where i have my global frame capital x capital y and capital z right this is the result of the translation of the global coordinate system the next thing that i am going to do is i am going to pan the camera this is i am basically doing the same way as i would imagine where the camera is being used
so If i have to pan the camera i will pan it until this x axis becomes parallel to the axis x axis of the image plane now in general given two coordinate systems you will not always be able to do it by a single rotation about one of the coordinate frame axis but in this case it will be possible because the way we have defined pan and tilt
so i am going to rotate this about the capital z axis until this coordinate axis capital x becomes parallel to small x
so what i will have now is my image plane in the same location with its x axis y axis z axis okay the center of the global coordinate system is still here what i have done is i have rotated about this vertical axis which is the z axis
so that does not change that remains where it was right what has happened is this x axis gone over there so now i have to draw a x axis which is parallel to the small x axis so i have capital x axis here and consequently this y axis is rotated so the y axis capital y axis is now here right you got that so we have rotated about the capital z axis from here

now the next thing i wanted to do is to tilt the camera so i take this now capital x axis which is parallel to small x axis and give it a tilt until this capital z becomes parallel to small z so now i have the situation where this is the camera coordinates center of the global coordinates are here x axis was already parallel to the two x axis were parallel to each other this z axis is now going to rotate until it becomes parallel to small z so this is the capital z axis obviously when these two are parallel the y axis is also going to become parallel

so this is my global coordinate with respect to the camera coordinates right finally i need to translate this from here to here the center of the global coordinate to coincide with the image coordinate

so that if you do you will get is my axis now they have become coincidental this is x and capital x this is y and capital y this is z and capital z and whatever my lens is if i take a point here i will be able to find a corresponding point so this this point here is my world homogeneous coordinates this is my camera homogeneous right so what this is if i go through the sequence having a translation okay having a rotation theta having a rotation alpha having a translation r then finally i have the perspective transformation or imaging transformation

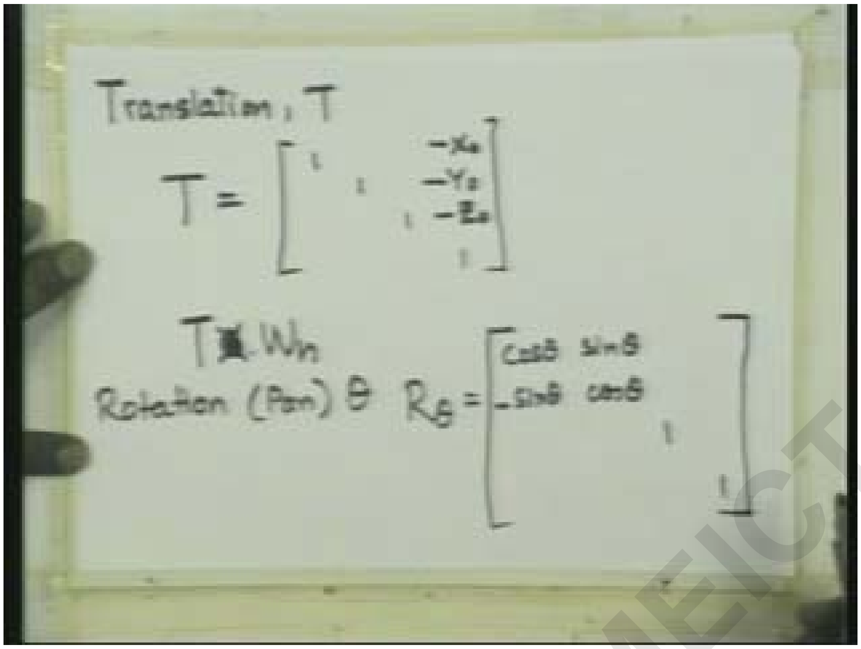
so if i do this part of the transformations i will be able to now be able to deal with any image point with respect to the global coordinates yes

we require because there is a option because unless you do it in that fashion it becomes very difficult to physically see it of course finally we can multiply all these transformations into one single transformation right

but how do we then physically interpret it becomes little more difficult so i have gone through this steps so that you will able to fully understand what we are talking about in terms of determining the final transformation

yes please yes yes [noise] thats what i said i said in this particular case it is possible because of the physical arrangement of the camera we are dealing with the situation here we are not talking about some arbitrary rotations of the camera

if you are talking about the arbitrary rotations of the camera then this logic will not apply strict we can find some other logic to get from one coordinates to the other coordinates but this particular logic will not work this simplifies to some extent but physically if the camera is mounted like this this is the only way it can move then you can always do that okay so now lets go and look at what are the transformations that we require okay so first translation (refer slide time45:44)



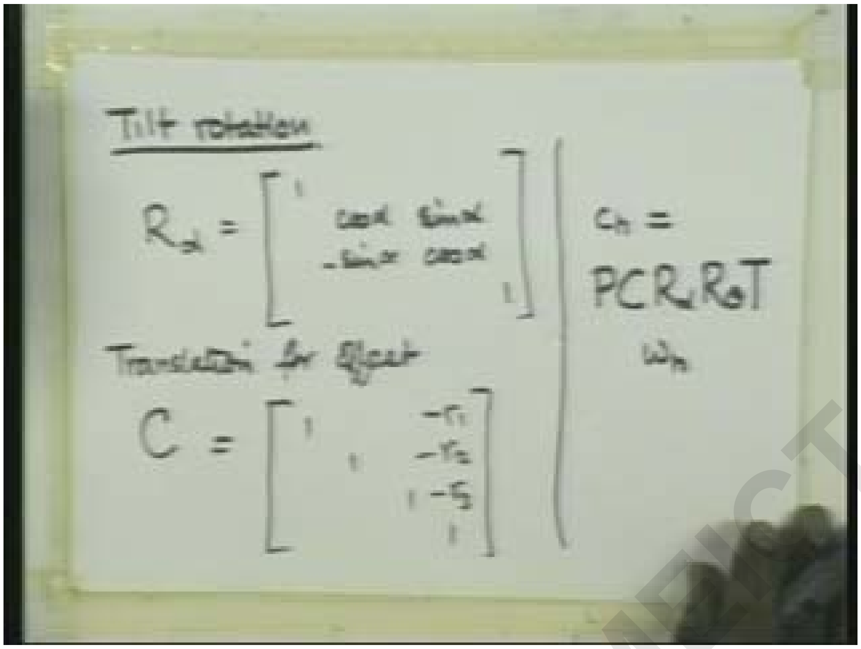
translation t what is the matrix which will be required for transformation capital t is one one one one minus x zero minus y zero minus z zero

go back here we are talking about this transformation t right what we are doing is we are translating this global x axis global coordinates to this particular point which is at the coordinates x not y not z not so for a given point say this is our world homogeneous or world point that we are interested in earlier it had some coordinates after this transformation in the global axis it is going to have some other coordinates

okay it is going to have coordinates which are going to be less by x not y not z not right because the coordinates of this particular point where we have translated the global x y z coordinates to this x not y not z not

so therefore this transformation matrix which i am showing you here is going to do the job because when you multiply this so If i multiply t by i called it w right

so okay so let me say this this is only multiplication a t times w t is this matrix w is the coordinates of that i will get the coordinates of the same point in the new coordinate plane thats what i wanted all right what about rotation the first rotation or the pan rotation or theta i will have a transformation matrix r theta which is going to look like what it will have i am rotating about the z axis so this is going to have cosine of theta sine of theta minus sine of theta cosine of theta one one i am deliberately not writing all the zeros in the matrix so that it does not get fluttered but all the other elements are zero right this you already studied in dealing with transformations let us look at the next translation next rotation which is the tilt rotation(refer slide time47:23)



okay the tilt rotation is r alpha which is now the rotation about x axis
so therefore i am going to have cosine of alpha sine of alpha minus sine of alpha cosine of alpha and unity
right then we have the translation so let me call that c translation for offset okay so that c is nothing but one one one one and you are going to have minus some coordinates r one, minus r two, minus r three
okay and therefore my final transformation is that i am looking for is camera coordinates equals to perspective transformation or imaging transformation times c times r alpha times r theta times t this whole thing when multiplied to w h will give me the camera coordinates okay
so systematically we have seen how we can do it you should look at it so that you are not confused about how the rotations are or how we have made the two coordinate systems coincident next time we will just touch up on it before we move on to the next subject thank you