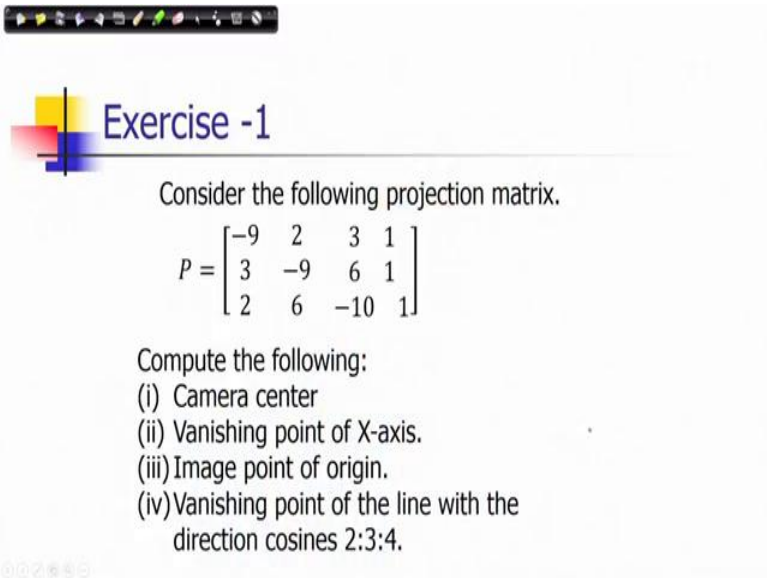
Computer Vision
Prof. Jayanta Mukhopadhyay
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

Lecture - 14
Camera Geometry Part - IV

We are discussing on single view Camera Geometry and we have discussed how different information can be obtained from a projection matrix. Now, let us discuss some of the exercises for your practice, solve some of the problems related to that it will give you a better insight.

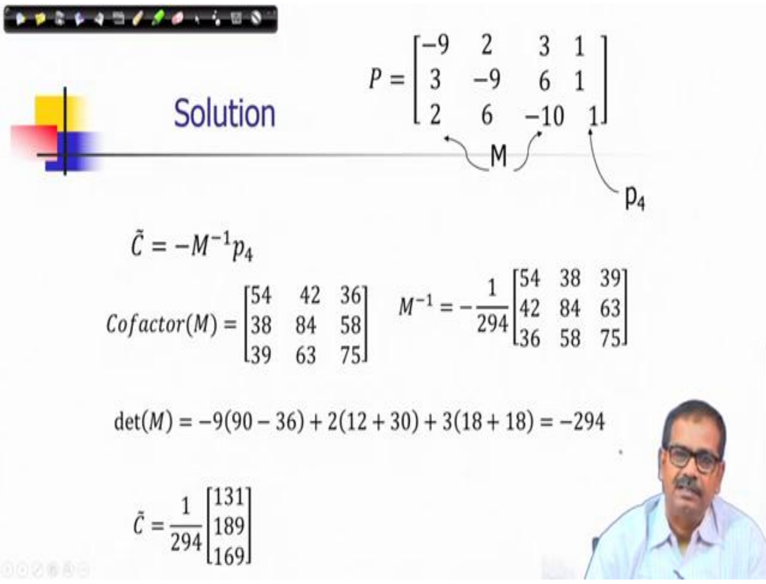
(Refer Slide Time: 00:44)



So, consider this problem that you know you have a projection matrix which is given in this form, and you need to compute the following, like you need to compute its camera center, vanishing point of X-axis, image point of origin, vanishing point of the line with the direction cosines say $2:3:4$

Now, here I would suggest that you may give a pause to my video lecture and solve this problem and then again resume it. So, I will discuss this particular solution in the next slide.

(Refer Slide Time: 01:18)

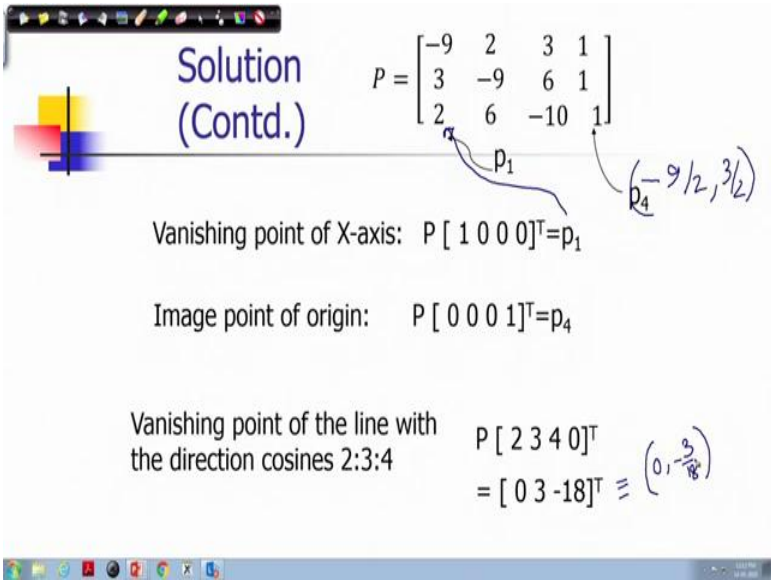


So, as I have shown that this is a projection matrix what you have considered. A 3×3 sub matrix M is defined in this form and then the other column vector that is a p_4 column vector. So, we can compute the camera center from this expression that we have to invert the matrix M and then compute $-M^{-1}p_4$ to get the camera center directly in the world coordinate system.

So, in this case if you carry on computations of inversion, the steps that you need to compute the cofactor of the matrix, you need to compute the determinant of the matrix and then the inverse can be computed in this form; that means, you have to take the transpose of the co-factor and then divide it by determinant. And then finally, if you perform these computations of $-M^{-1}p_4$ you will get $\tilde{C}$ in this form.

So, you should note that this is just in the world coordinate system itself you need not convert, this you should not convert it into a non-homogeneous coordinate system then it will be a wrong. So, it is directly it is giving you the result in a three-dimensional coordinate system as we understand.

(Refer Slide Time: 02:37)



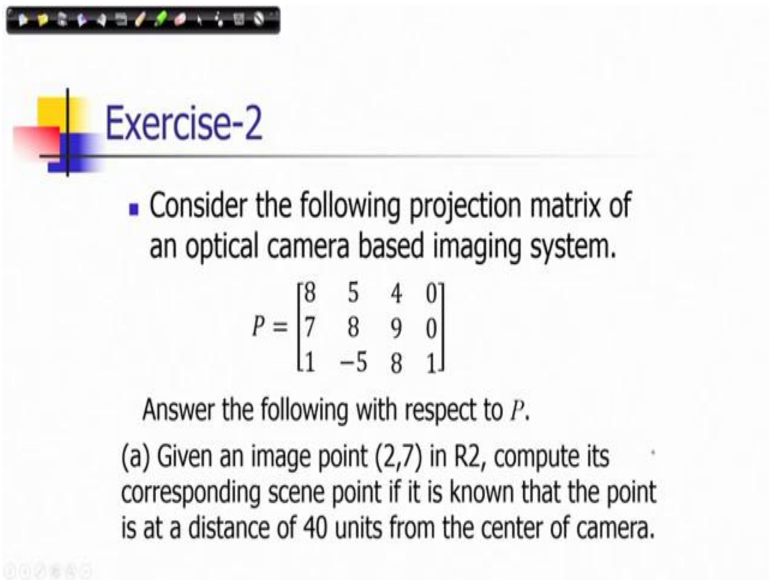
Let us considered how do you compute the vanishing point of X-axis. As I discussed earlier that for vanishing point of X-axis we have to multiply the projection matrix with the directions along X-axis and the additional scale value, scale dimensional is 0 because that represents a point at infinity along X-axis. So, you have to multiply projection matrix with this vector. I am showing for the gravity in the row form. It is a column vector $\begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}^T$ and then you will get p_1 and which means a you will get particularly you will get $\begin{bmatrix}-9 & 3 & 2\end{bmatrix}^T$.

So, this is related with $\begin{bmatrix}-9 & 3 & 2\end{bmatrix}^T$ that is a column vector. And you should note that this is actually in the homogeneous coordinate system in the image coordinate system. So, if I would like to get them in our two-dimensional coordinate system, non-homogeneous coordinate system I have to express the x coordinate as $\frac{-9}{2}$ and y coordinate as $\frac{3}{2}$. So, in a image coordinates in our understanding of normal two-dimensional coordinate system is this point is at $(-9/2, \quad 3/2)$.

Similarly, for image point of origin that is the 4th column vector, so you will get $\begin{bmatrix}1\\1\\1\end{bmatrix}$.

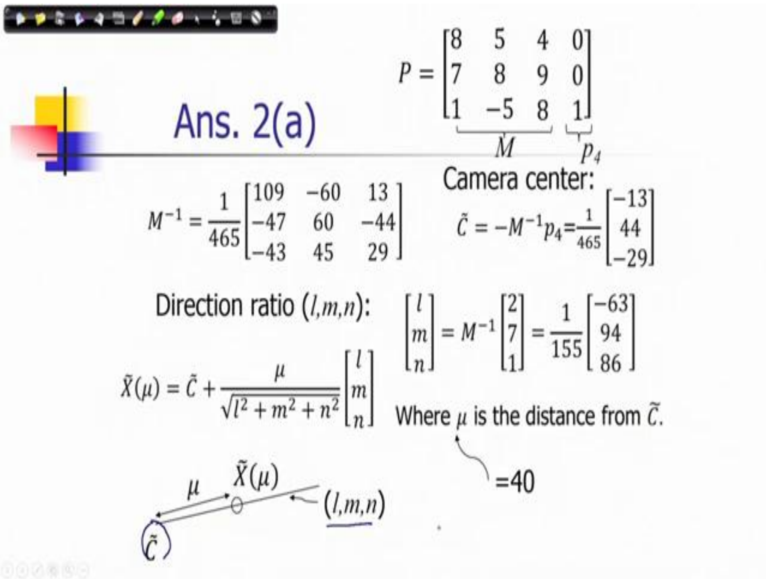
So, if I adjust the scale it is the coordinate 1 and 1 1. And then the vanishing point of the line with the direction cosines $2:3:4$. Once again this shows that, what is the point at infinity along the directions which is expressed by the vector . $[2 \quad 3 \quad 4 \quad 0]^T$ and if I multiplied with projection matrix P. So, you will get a vector $[0 \quad 3 \quad -18]$. Once again this is in the homogeneous coordinate system which means this is equivalent to a coordinate $(0, \quad -3/18)$ in our two-dimensional real space.

(Refer Slide Time: 04:50)



So, let us considered another example; here also you considered a projection matrix given in this form and then you are asked to do this computation that if you have an image point (2, 7) in R2 or in R square that is normal to dimensional coordinate convention you have to compute its corresponding scene point if it is known that the point is at a distance of 40 units from the center of camera. So, once again I would request you to pause the video at this point, solve this problem and then again resume it. I am going to discuss this solution in the next slide.

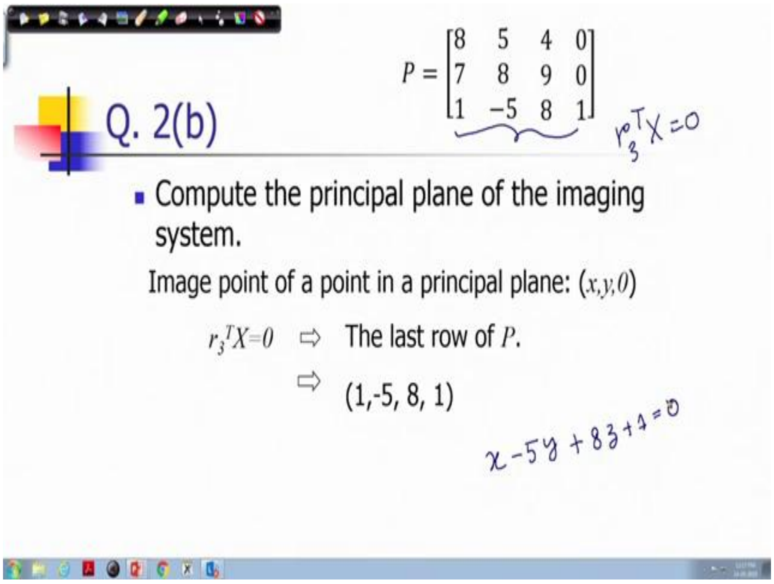
(Refer Slide Time: 05:33)



So, here I have shown you some of the structures of this solution. As you can see that projection matrix is represented by those sub matrixes M and p_4, M is a 3×3 sub matrix p_4 is that column vector. So, camera center first you need to compute because once you have to compute the corresponding projection ray. What we required? We required the camera center $\widetilde{C}$. So, require the camera center $\widetilde{C}$ and we required the direction cosine (l,m,n). So, $\widetilde{C}$ is given by this ($\widetilde{C} = -M^{-1}p_4$) relation. So, you have to compute M^{-1} and then $-M^{-1}p_4$ will give you corresponding $\widetilde{C}$. It is again in the three-dimensional coordinate space whereas, the direction cosine should be given by $M^{-1}\begin{bmatrix} 2 \\ 7 \\ 1 \end{bmatrix}$.

So, (2, 7) was the image point and along that, directions in the homogenous system coordinate system it is (2, 7, 1). So, if I perform these computations, I will be getting this particular direction (l,m,n). And then any point in this ray which is lying at a distance μ can be expressed in this form which is $\widetilde{X}(\mu) = \widetilde{C} + \frac{\mu}{\sqrt{l^2+m^2+n^2}}\begin{bmatrix} l \\ m \\ n \end{bmatrix}$. So, you see that this is nothing, but the unit vector along the direction (l,m,n). So, we can always express in a parametric form equation of this straight line in this form. So, that is how you can get the corresponding projection ray. You can put $\mu = 40$ and that would give you the corresponding ray. So, this is the solution of this point.

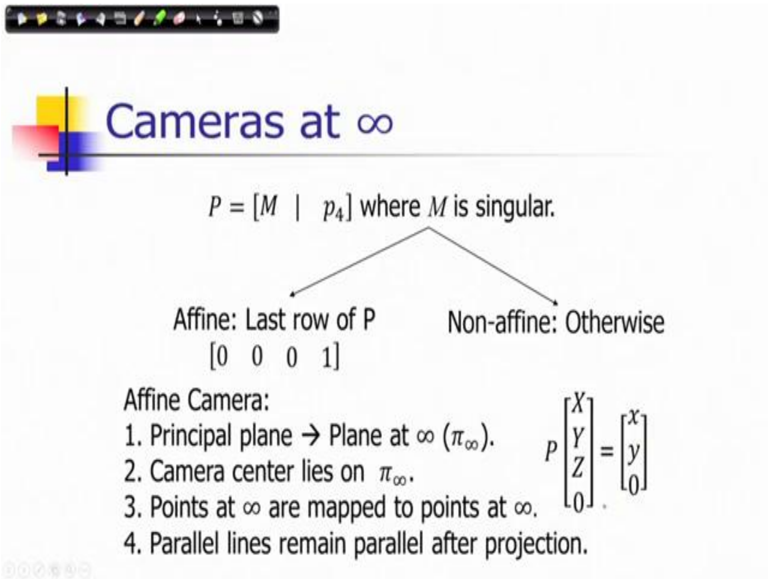
(Refer Slide Time: 07:23)



We continue the same problem. Here you are asked to compute the principal plane of the imaging system with the projection matrix. So, if I start doing this operation how do you get the principal plane. It is as we know this is the principal plane is given by this third row. So, which is given by $r_3^T X = 0$.

So, you will find that this solution is basically it is an explanation one again. Image point of a point in a principal plane would be in this form, so it should be $r_3^T X = 0$. So, it is a last row of P and it is given by (1,-5,8,1). If I write in our normal convention of planner equation it is $x - 5y + 8z + 1 = 0$. So, this is how you get the principal point.

(Refer Slide Time: 08:42)



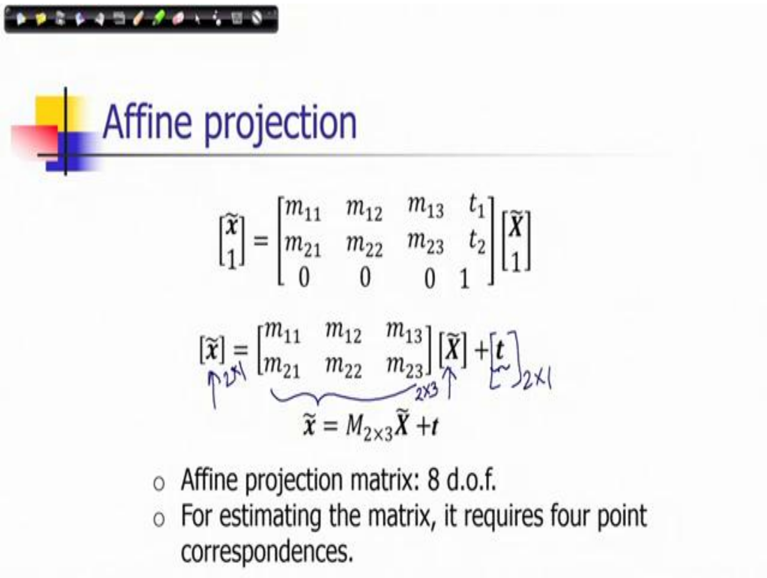
So, now we will continue our discussions in this topic. So, let us consider another scenario where camera center is at infinity. Now, in this case as we have discussed earlier also that M would be singular that is the first thing, because in that case the 0 of M will give you the center and the third, fourth dimension of that center that is a scale factor would be 0. Now, there are two situations in this case it could be an affine camera or it could be non-affine camera. So, we are not interested on non-affine cameras for this kind of situation we will be consider in only affine camera.

So, one of the simple characteristics of affine camera is that its last row is in the form of [0 0 0 1], you can use any other scale factor instead of one non-zero value, but canonically you can put it as [0 0 0 1]. Then every other element gets fixed with respect to that scale 1. So, the property of affine camera is that first its principal plane is a plane at infinity that is a principal plane; that means, all the points which are lying at infinity in the along many directions. They are lying at in the principal plane.

And it camera center also lies on the principal plane naturally. So, point set infinity are map to point set infinity. So, if you consider any points which are at infinity, its vanishing points, so called vanishing point; that means, if you take the image of that point that would be also at infinity; that means, it is scale value would be also 0 which means that as you understand that intersection of parallel lines of a non homogeneous coordinate at parallel lines in our Euclidean geometric sense can be explained as intersection at a point at infinity which is represented in this homogeneous coordinate system where scale value is 0.

So, still the lines after transformation, which means after getting imaged still they remain parallel. because images of those lines are also their intersection is also at infinity. So, that is one of you know major property of this particular things and it explains why a points at infinity, still map to a point at infinity.

(Refer Slide Time: 11:28)



So, the affine projection one of the simplification of the camera relationship can be done in this way as you can see that, in this form in affine camera this is the world coordinate system in the non-homogeneous coordinate system which means it is simply a three-dimensional point. You are multiplying this three-dimensional point by a matrix in the form of 2×3 and then you will get another vector three-dimensional vector and this is a parameter which we called as a translation parameter for this affine transformation and then you get the corresponding image point. So, this is a image point of $\tilde{X}$.

So, that is also in the non homogeneous coordinate system because this will give you 2×1. So, this is actual image coordinate what will get you do not have to do any scale adjustment. So, this is a very simple relationships in affine geometry where you can express the relationships all in the sense of Euclidean space, Euclidean space this relationships in the sense of three-dimensional Euclidean space to a two-dimensional Euclidean space of the images.

Now, this can be explained how do you get this kind of structure, if I considered the canonical representation of the affine projection matrix where you can see that the last row of this projection matrix is given by [0 0 0 1] vector. So, if I do the matrix multiplication and consider the corresponding sub matrix multiplication then it is equivalently coming in to this point. So, this simplifies the relations between the projection, the image point with the world coordinate point and it simplifies also to the computation of finding out this affine matrix.

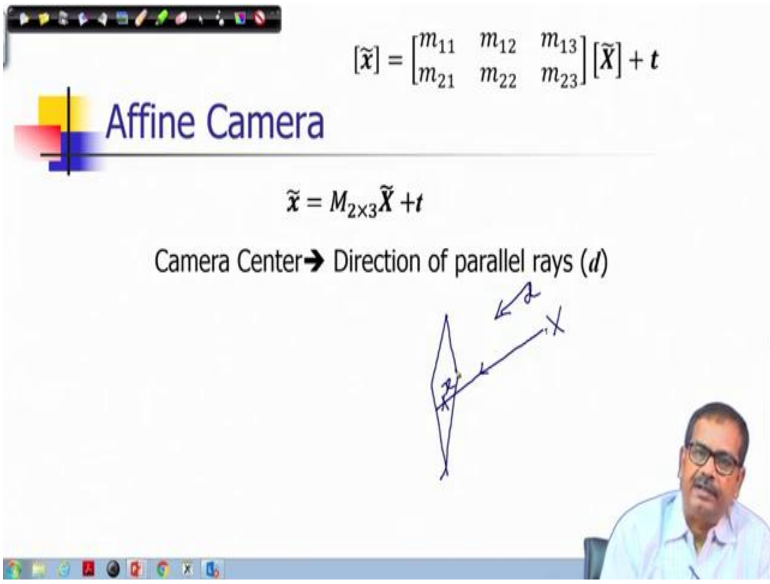
So, let us see how you can do it. So, you can express this equation in the following form

$$\tilde{x} = M_{2\times3}\tilde{X} + t$$

$M_{2\times3}$ is a sub matrix that is a matrix represented above and $\tilde{X}$ represents the same point plus t. So, affine projection, so how many independent parameters are there? As we can see that this t is 2×1 vector and this(M) is 2×3 so there are 8 parameters. So, there are 8 independent parameters or 8 degrees of freedom.

So, this is the final conclusion about or this is a advantage what you have, you required less number of points to estimate this particular projection matrix because you have only 8 independent parameters it requires 4 point correspondences each point correspondence is giving you two equations, one for x coordinate another for y coordinate. So, if I have 4 point correspondences I will get 8 equations and I can solve this problem. So, that is the minimum requirement. If you have more number of point correspondences then you can perform least square estimate, where which we will discuss in the next slide.

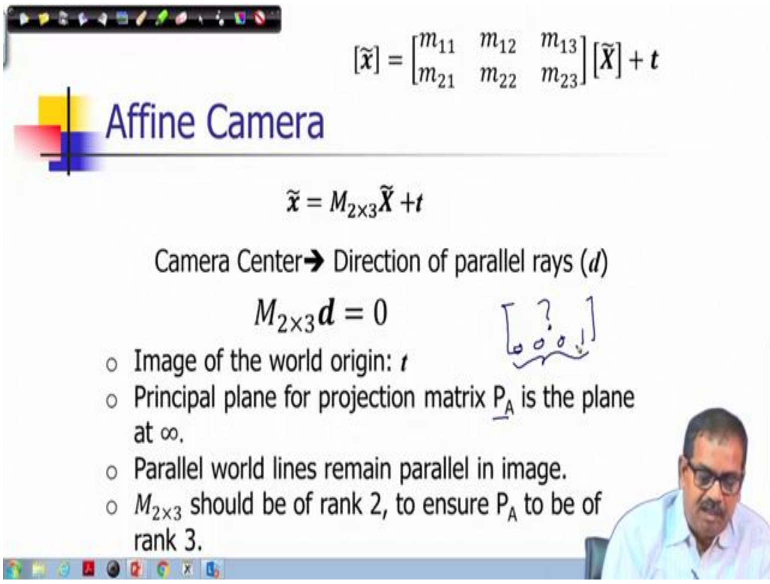
(Refer Slide Time: 14:58)



So, this is how will be discussing. So, first thing is that in the affine camera center it lies at infinity and it is a direction of parallel rays. So, geometric interpretation is that: in an affine camera the imaging takes place using parallel rays, instead of considering a particular center where all the rays are connecting to that center that is a perspective projection geometry. In affine geometry you considered any direction any parallel to any direction and any particular direction, let me explain.

Suppose this is your image plane and suppose this is your scene point X and say the rule is that this is a direction of the vector d; so, all lines parallel to the. So, X, image of X would be that you draw a parallel line along this direction and where it intersects that becomes a image of this point. So, this is a interpretation of imaging for affine camera which is the parallel projection in the case of imaging geometry.

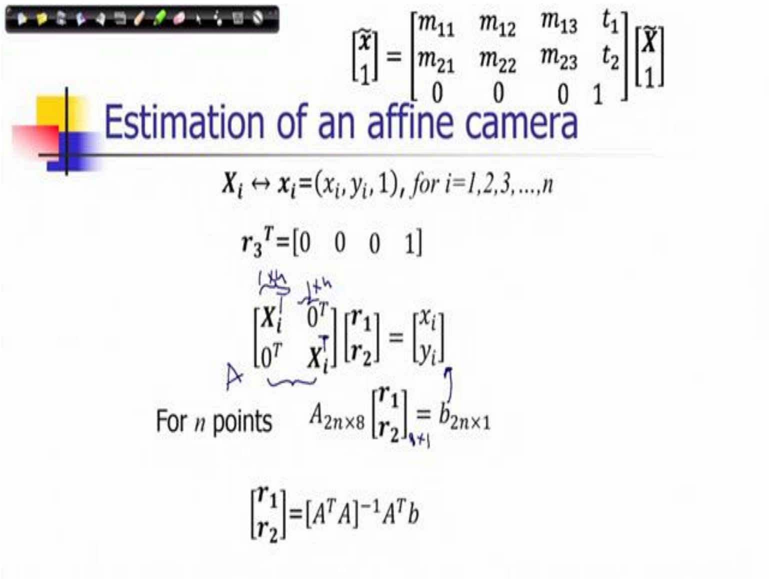
(Refer Slide Time: 16:10)



So, this is the relationship with M and d, you can get d in this by exploiting this equation $M_{2\times3}d = 0$. So, the interpretation of t that it is the image of the world origin and its principal plane or affine projection matrix is a plane at infinity that you can see that [0 0 0 1]. So, that is the form.

So, in a in P_A the last row is [0 0 0 1], in the last row of the affine matrix. So, this denotes the principal plane and that is the plane at infinity that is a interpretation in the projective space and $M_{2\times3}$ should be of rank 2, to ensure that P_A to be of rank 3. So, these are certain interpretations.

(Refer Slide Time: 17:11)



So, let us considered estimation of an affine camera when you have more number of points; you require minimum 4 points, but if you have more number of point correspondences. So, you can minimum 4 point correspondences and if you have more number of point correspondences then your estimation would be robust. So, this equation so, this is one example has been shown specifications that capital X_i and x_i they are the point correspondences X_i is the scene point and corresponding image point is x_i shown in bold font. So, in the non-bold font its coordinates are expressed; so, these equations if you form this equation.

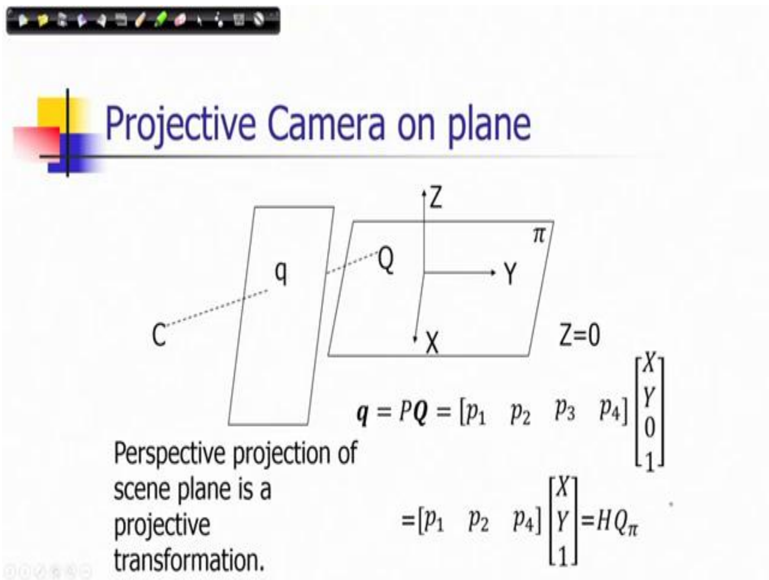
So, once again I will be considering the concatenation of row vectors as the parameters of my projection matrix. So, this could be represented by these two equations. So, you can see that in this case $X_i r_1$ is giving you x_i, x_i coordinate and $X_i r_2$ is giving in the y_i coordinate. So, x_i should be represented in the form of a transpose here which is not shown here, mathematically it should be X_i^T transpose and there should be also X_i^T transpose because you know they are row vector and then you can form this equations.

So, for n points, this matrix I can consider this matrix ($\begin{bmatrix}X_i^T & 0^T\\ 0^T & X_i^T\end{bmatrix}$) as matrix A and this could be considered a matrix say A. And if there are n points so, each one will give me two equations. So, there will be 2n equations. And the dimensions as you can see this is this is 1×4 and this is 1×4, so it will make it 8. So, it is $2n \times 8$ and this dimension is also this is 8×1. So, you get $2n \times 1$. And this matrix this column vector is found by the coordinates respective coordinates. So, this is the set of equations which you need to solve and this is a non-homogenous set of equations because you do not expect that every coordinate would be 0. So, this vector is not going to be a 0 in your experimentations or observations.

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = [A^T A]^{-1} A^T b$$

So, you can solve it by using standard least square error method for non-homogeneous set of equation. And the solution is given in this form I discussed the nature of solution, so I can use a $[A^T A]^{-1} A^T b$ that would give me in the solution.

(Refer Slide Time: 20:04)



So, we will discuss this particular thing that how the points which are lying on a plane form images using projective camera. Without loss of generality let us consider that plane is the say XY plane where Z is equal to 0. You can always make coordinate transformation to make any plane as XY plane and apply this principal that is why I said it is without loss of generality.

And say a point is given as q that is a scene point Q and this is a camera configuration, where C is the center of projection and there is a image plane. So, a ray formed between

C and Q and the intersection of that ray with the image plane which is shown by the point l q that is image point. So, this is how the imaging takes place.

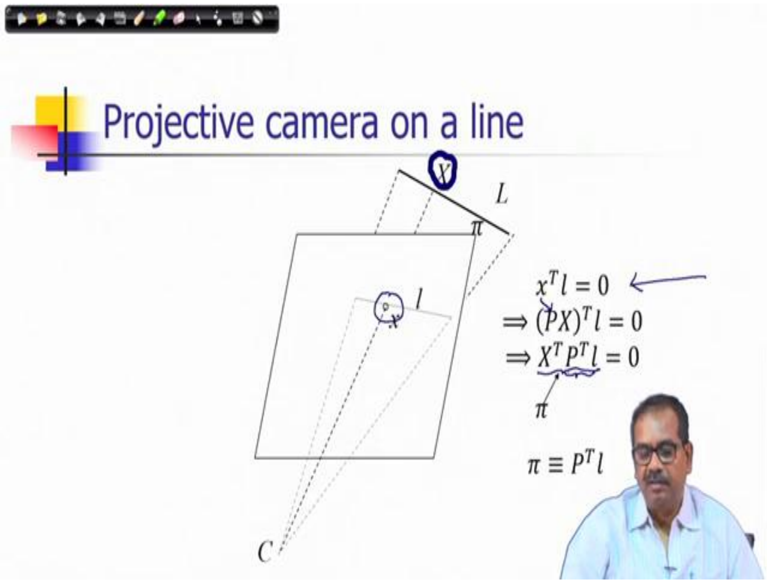
So, now, I can express the coordinate q in this way.

$$q = PQ = [p_1 \quad p_2 \quad p_3 \quad p_4]\begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix}$$

You can see here that any point in that XY plane can be written as it is coordinate as $\begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix}$. So, finally, this relationship in this projection geometry projection matrix and multiplication projection matrix with the three-dimensional point in the projection in the homogenous coordinate system can be reduced to a form where you require only two coordinates X Y which is a point in that plane. So, point in the coordinate convention of the plane and multiplied by 3×3 matrix.

We know this form that this is nothing, but a 3×3 matrix transformation of a homogenous point in a two-dimension, projectives two-dimensional projective space to another two-dimensional projective space which is like in homographic. So, if you take the imaging of a plane it establishes homography between the image points and the scene points, that is the crux of this discussion. So, that is what? Perspective projection of same plane is a projective transformation.

(Refer Slide Time: 22:27)



We will discuss on imaging of a three-dimensional straight line and try to see how the corresponding projected line on the image plane is related to the three-dimensional configuration. So, consider a line L that is a three-dimensional line as shown in the diagram and you have an image plane and its camera center at C. So, if you would like to project this line, I can consider any two end point and of this line any two points in on the line and find out its image points and connect those image points.

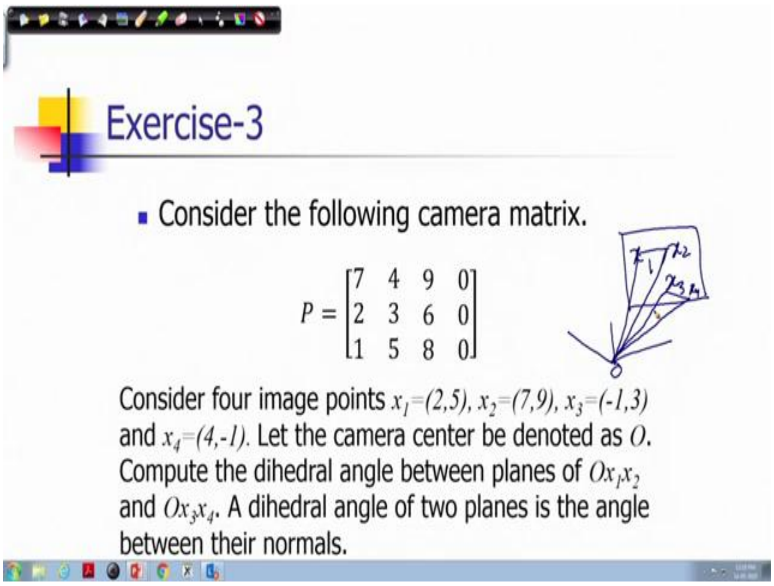
You will get another line that is a line of that image that would be a straight line also because this property of projective transformation that is applied in this case. Now, let this line is denoted by l in my figure and as you understand this line is represented in a two-dimensional projective space. So, what kind of three-dimensional information that we can recover if I know the projection matrix, that I would like to discuss here.

So, considered this that there is a point on a three-dimensional straight line and which is given in this form that this is X this straight line and it is corresponding image point is shown here by drawing the projected ray that is x. So, the relationship between x and l, that can be found from the point containment relation. And you can also see that it forms a plane, a three-dimensional plane which connects the camera center and also three-dimensional straight line.

Note that the image of that straight line that also lies on that plane. So, now, as I was mentioning that a point containment relation of the image point x can be expressed here that is $x^T l = 0$. And if I reduce x if I express x in terms of imaging of a three-dimensional points, so that is you have to multiply the three-dimensional point in its homogeneous representation with a camera matrix P. So, you get PX equals that image point x.

So, $(PX)^T l = 0$ and using the matrix property I can convert this expression as $X^T P^T l = 0$. So, note that this relationship is again a relationship of point containment in a plane where the plane is given by this $P^T l$ So, we can get the expression of the plane by this computation give in the projection matrix we can and also given a straight line we can find out the plane on which the straight line camera center and their image line.

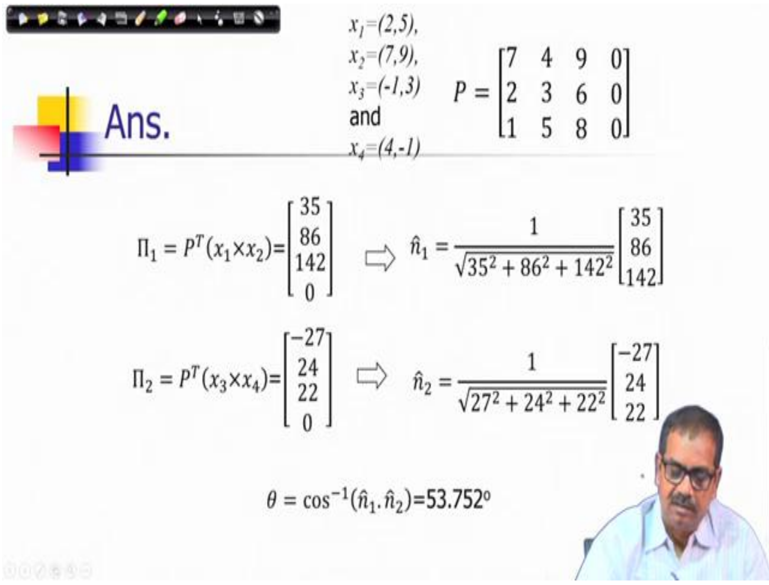
(Refer Slide Time: 26:03)



The third exercise here you have this projection matrix and considered 4 image points x_1, x_2, x_3 and x_4 given in this form and you denote the camera center as the origin as a point O, it is not origin it is a point O and you have to compute the dihedral angle between planes of Ox_1x_2 and Ox_3x_4 . And how a dihedral angle is defined? It is a angle of two planes, it is a angle between their normal. So, angle between the normals of those two planes.

So, this problem you need to solve. Just to show you diagrammatically what I have asked you that consider an image plane and you have points say x_1, x_2 and say x_3, x_4 . So, and say this is the camera center; this is a camera center O. So, you can form a plane. So, 3 points can define a plane. Similarly, you can form another plane, now what is the angle between these two plane; that is the problem what you need to solve.

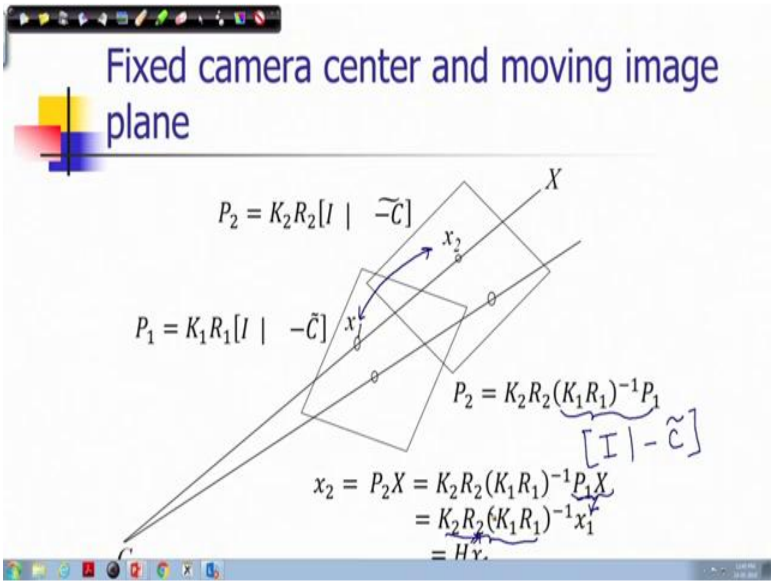
(Refer Slide Time: 27:37)



So, this is how the solution we will get that once again this is summary that those are the points shown and also the projection matrix shown here. So, first you have to form the first dihedral plane. So, you can see that by performing the cross product of x_1 and x_2 you are getting the line found by x_1 and x_2, and we know that given a line in the image plane how the three-dimensional plane can be obtained. It is P transpose in to the line in its corresponding image coordinate system itself, in the image representation itself. ($P^T(x_1 \times x_2)$)

So, if I perform this operation then I would get the first plane as given here $\begin{bmatrix} 35 \\ 86 \\ 142 \\ 0 \end{bmatrix}$. So, this is a first plane. Similarly, for the second plane which is found by the images image point x_3, x_4 and the center of camera O that is also given similarly in this form. So, now, these are the two planes. If you take its corresponding normal vectors unit normal vectors you can compute in this form for the first plane, for the second plane. Now, to compute the angle between this two you can take the dot product of this unit vector and applied that cosine law. So, cos inverse of this will give you that angle; so, this is the answer. It is coming around 53.752 degree in this case.

(Refer Slide Time: 29:09)



Next will discuss about the relationship between the image points when your camera center is fixed and its image plane is changing. So, you consider this two scenarios that there are two image shorts, one of them is taken in this configuration when this is the image plane and this x_1 is the image of the scene point x. As you can see camera center is still remaining at this position.

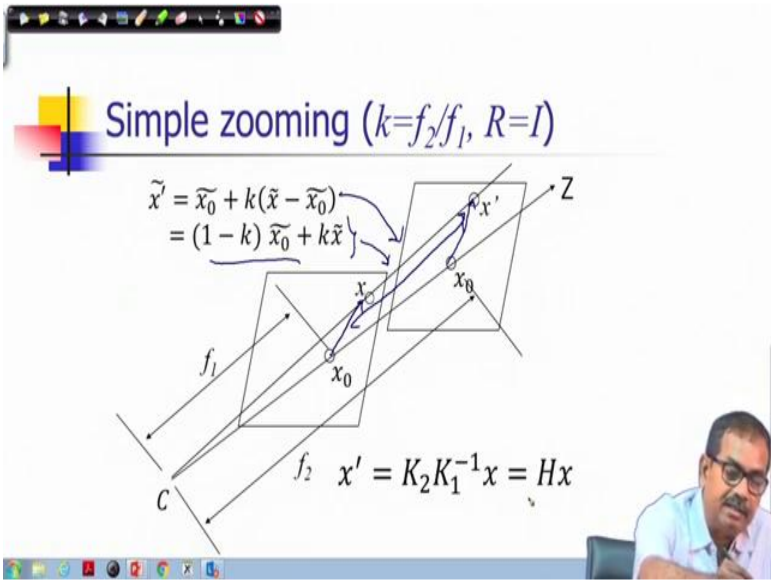
So, the camera projection matrix is given by P_1 in this case and which is represented in the general form in this form($P_1 = K_1R_1[I \mid -\tilde{C}]$) using this camera center, where K_1 is a calibration matrix and R_1 is a rotation matrix and the other for the other imaging system. So, this is the image plane and its corresponding projection matrix is P_2 which is given in this form($P_2 = K_2R_2[I \mid -\tilde{C}]$). We can relate P_2 and P_1 in this way, we can see that this P_1 if I apply $(K_1R_1)^{-1}P_1$, this is giving you this $[I \mid -\tilde{C}]$. So, this is giving you this particular expression. So, finally, if I multiplied with K_2R_2 then it becomes P_2 that is how this relationship is established.

So, if I apply this relationship then we can find out that this two image point they form a homography that we discussed already in our two-dimensional, discussion on two-dimensional projective geometry and projective transformation in particular. So, I can express this homography in this form. I will follow this. So, I will start from x_2 which is given by the image of X in the camera P_2. So, P_2X.

$$x_2 = P_2X = K_2R_2(K_1R_1)^{-1}P_1X = K_2R_2(K_1R_1)^{-1}x_1 = Hx_1$$

Now, P_2 can be related with P_1 in this form and then as you can see this P_1X can be reduced to the scene point of the other camera x_1 for the same corresponding scene point. And then the relationship between x_2 and x_1 can be given as Hx_1, where H is equal to this ($K_2R_2(K_1R_1)^{-1}$) matrix and it should be invertible. So, this is also expressing the projective transformation between the corresponding image points. So, from the camera geometry itself we can explain this relation.

(Refer Slide Time: 32:05)



And when we considered the image planes their parallel then the phenomena can be expressed like a zooming of images. So, image planes they are parallel, as if you are changing the focal length; that means, the direct that is a distance from the camera center to the image plane. So, you are changing the focal length. And let us consider the ratios of this focal length is k, and we know that under this situation the homography is established because there is no rotation in between these two cameras.

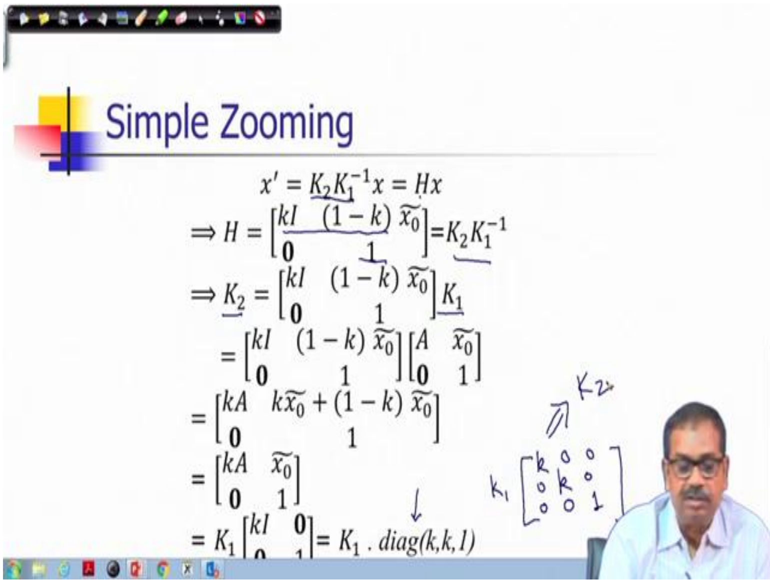
So, if the rotation is identity matrix then in a simpler form this relationship can be expressed as $K_2 K_1^{-1} x$. $K_2 K_1^{-1}$ will give you the corresponding homography matrix. So, this is the f_2 that is the distance from center and this is f_1 and K is a ratio.

Now, this can be explained in a further. we can elaborate this relationship like if I considered the deviation from any point from a particular, no. So, x naught is a principal point. So, if I consider deviation this vector from the principle point now this deviations or this distances or this vectors it should be also scaled in the same amount. So, it is a vector direction remains same, but it is scaled by this factor k.

$$\tilde{x}_0^{'} = \tilde{x}_0 + k(\tilde{x} - \tilde{x}_0) = (1-k)\tilde{x}_0 + k\tilde{x}$$

So, any point can be expressed in this form given in this equation. So, $\tilde{x}_0$ and it is called in its normal two-dimensional coordinate system, k in to this vector that would give you corresponding from here, from this position the from the geometry itself you can use this relation and this is a expression. Now, this expression can show the structure of this homography matrix. How it is so?

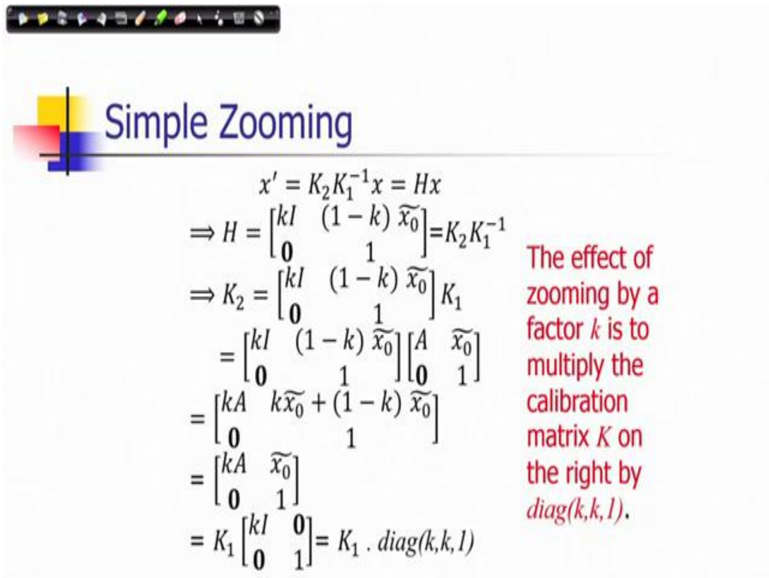
(Refer Slide Time: 34:23)



So, let me show you in the next slide. So, this is the structure as we have mentioned that if you note that your homography structure is $K_2 K_1^{-1}$ and finally, using that relationship using that deviations I can relate H with this elements $\begin{bmatrix} kI & (1-k)\tilde{x}_0 \\ 0 & 1 \end{bmatrix}$. So, this is equal to $K_2 K_1^{-1}$. And from using this relationship I can relate the calibration matrices between the two configurations K_2 and K_1. So, what we will observed that actually the calibration matrix is nothing, but K_2 calibration matrix is nothing but related with K_1 by simply multiplying by a diagonal matrix k k 1 (diag(k,k,1)).

I already explain how this diagonal notation should be interpreted. You should have a diagonal elements as k k 1, these are the 0s ($\begin{bmatrix} k & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{bmatrix}$). So, if I multiply K_1 with this matrix then I will get K_2. So, this is the relationships when you have zooming.

(Refer Slide Time: 35:35)



So, the effect of zooming is that you had simply multiplying the calibration matrix on the right by diagonal. With this let me stop here for this lecture. We will continue in the next part. We will still continue this topic where in our next lecture.

Thank you very much for listening.