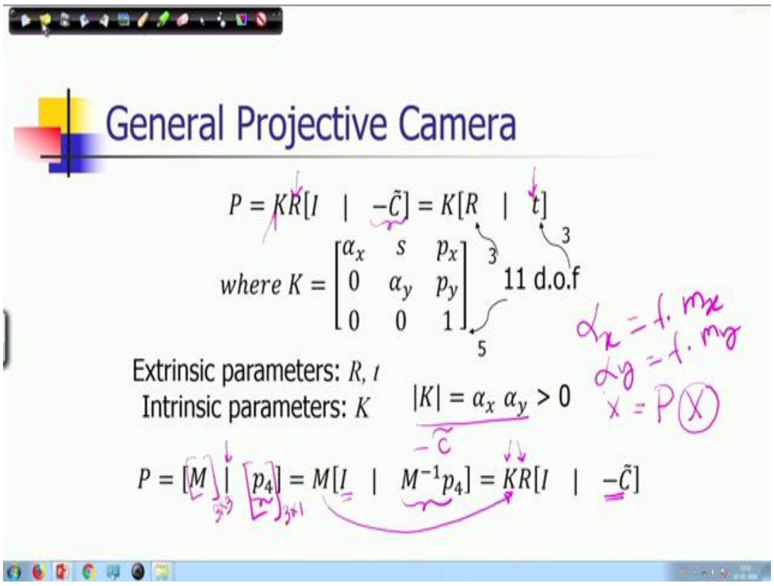
Computer Vision
Prof. Jayanta Mukhopadhyay
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

Lecture - 12
Camera Geometry Part - II

(Refer Slide Time: 00:21)



We will continue our discussion on Camera Geometry. In the previous lecture, we discussed the form of a general projective camera. So, there is a projection matrix which maps a 3 dimensional coordinate point to a 2 dimensional image point and in the projective space, we can represent them as in this form.

$P = [M \mid p_4] = M[I \mid M^{-1}p_4] = KR[I \mid -\tilde{C}]$ are the different kinds of representation of the projection matrix.

So, we can say that there is a 3 dimensional point X and if I multiply with the projection matrix then we will get a 2 dimensional image point $x = PX$ and in the homogenous coordinate system x is a 3 vector and X is a 4 vector and we can see that a projection matrix has different components here it has the calibration matrix and then there are different forms of representation of this projection matrix.

$P = KR[I \mid -\tilde{C}] = K[R \mid t]$

It can be shown as a combination of different types of matrices like in the above representation. it is K into R whis is the rotation matrix R which is involved coordinate transformation from all coordinate to the camera centric coordinate by rotating the coordinate access, it gets aligned with the camera centric coordinate system. And then the translation of the origin that is also represented by this parameter and in fact, this is the centre of the camera centre this is also shown here. So, these are the different kinds of information which is embedded in this particular matrix representation.

You should note the feature of a camera calibration matrix, it is an upper triangular matrix and if I take the determinant of this camera calibration matrix this should be the product of those two resolution factors α_x, α_y they are expressed in terms of number of pixels for representing the focal lengths. they are involved with the resolutions along x directions and y directions.

So, α_x, α_y are involved $\alpha_x = fm_x$ where f is a focal length and m_x is the number of pixels along horizontal direction and $\alpha_y = fm_y$ where, my is a number of pixels along vertical directions and f is focal length. So, this is what we discussed in the previous lecture also you should note there are different ways of notations by which we will refer this projection matrix say a projection matrix is a 3×4 matrix.

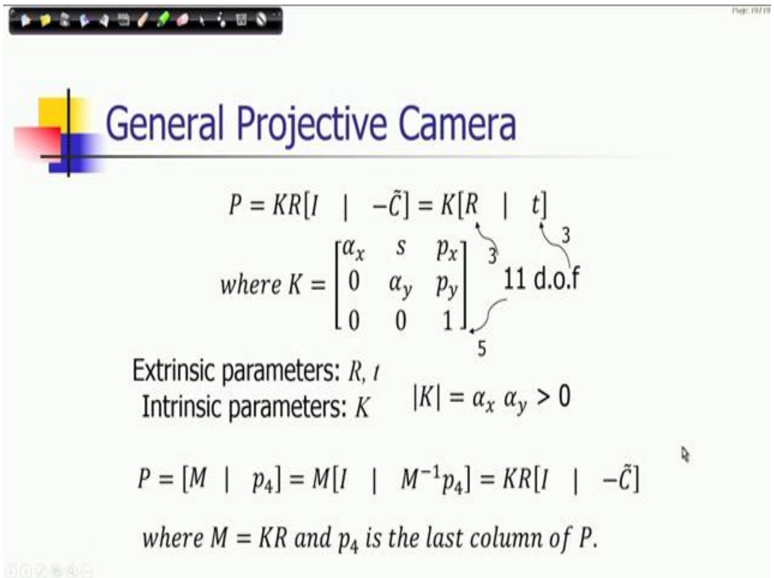
So, M out of them, the first 3×3 sub matrix that left side of sub matrix with 3×3 sub matrix is denoted by a symbol say M and the fourth column vector; that means, this is a 3×1 column vector that is also denoted by p_4.

$$P = [M \mid p_4] = M[I \mid M^{-1}p_4] = KR[I \mid -\widetilde{C}]$$

So, this representation is M and this separation just show that this is a sub matrix and this is p_4. Another representation if I take M outside then I can consider this is an identity matrix which is a kind of canonical form of representation of projection matrix if you remember that is the identity part of the sub matrix and this is nothing, but the negation of the centre; that means, the this $M^{-1}p_4$ gives you the camera centre negation of the camera centre.

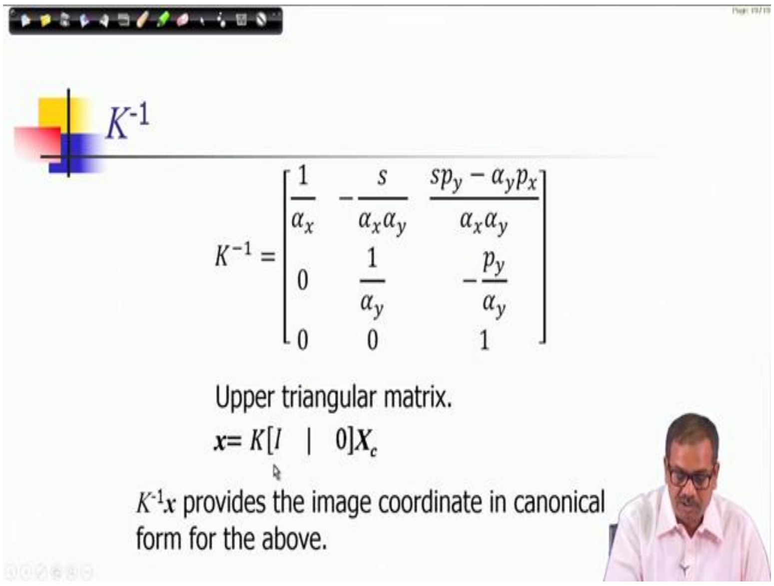
So, we can also represent this element as $-\tilde{C}$ and M itself can be decomposed into 2 3×3 matrix, it has a component calibration matrix, the other component is a rotation matrix. So, these are different ways this projection matrices can be represented and it could be understood all in those forms.

(Refer Slide Time: 04:49)



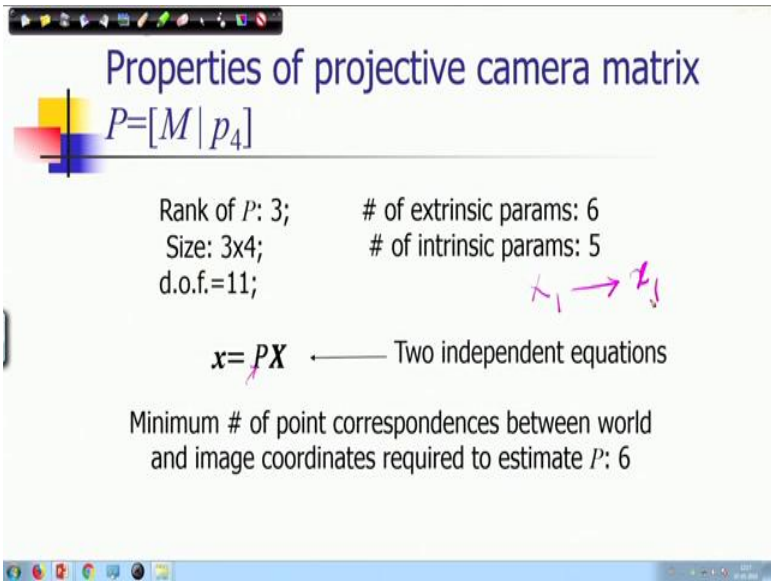
So, this is what I mentioned M is the product of K and R and p_4 is the last column of the projection matrix P.

(Refer Slide Time: 04:57)



Another thing we should note about the inverse of this calibration matrix, this is also interesting to note that this calibration matrix is also an upper triangular matrix and since the relationship between the image coordinate point and the corresponding camera centric coordinate system is in this form. So, if I apply the $K^{-1}x$; that means, if I multiply the image coordinate with this inverse of calibration matrix, it will provide you the corresponding coordinate in the canonical form which means the image coordinates in the canonical form where the focal plane is at the unit distance and its axis are parallel to the axis of the camera centric coordinate system. So, this is our interpretation.

(Refer Slide Time: 05:51)



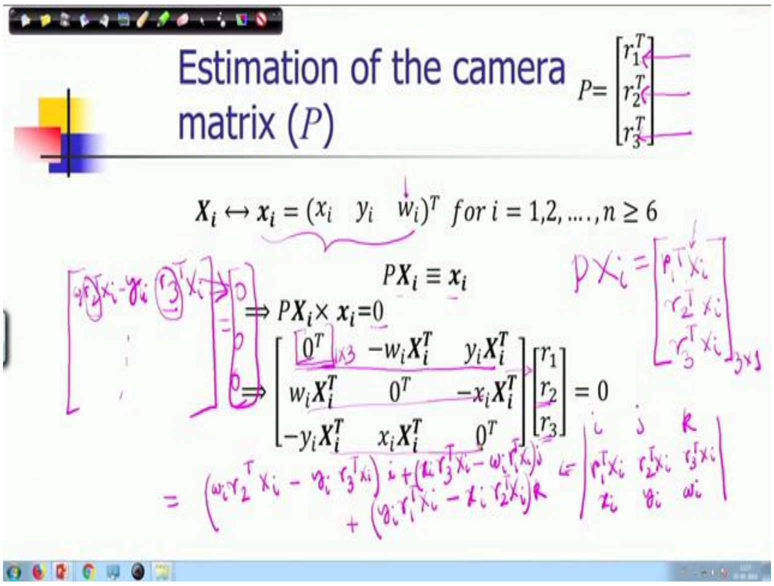
So, we will summarize here. we will look at some of the properties of the projective camera matrix. So, first thing as I mentioned the interpretation of the projection matrix is that if I multiply it with the 3 dimensional coordinate point, you will get the 2 dimensional image point and the rank of this projection matrix is 3, it is a 3×4 matrix, its size is 3×4 and if the number of independent parameters in the projection matrix is 11. So, degree of freedom is 11.

So, out of which number of extrinsic parameters that is 6 that we have discussed, 3 rotation parameters those are involved in forming rotation matrix; that means, 3 rotations of axis and 3 parameters for translation of origin and number of intrinsic parameters that is 5 as we have already seen the in the calibration matrix it is an upper triangular matrix and there are 5 independent parameters. And when we express this relationship when we expand this relationships in terms of coordinates individual coordinates, we will get actually two independent equations.

So, we will see how those equations can be written in the next slides. So, since there are two independent equations; that means, if I give you certain point correspondences suppose, the problem here is that I give you the scene points say X_1 and also its corresponding you know image point small x_1. So, if it is given to you then I can apply this equations, I will get two equations, but how many unknowns are there as I mentioned in the projection matrix there are 11 unknowns. So, I should get at least 11 equations to solve for all the unknowns, but since 1 point correspondence gives me 2 equations.

So, I require at least 6 such point correspondences, I require at least 6 such point correspondences to form you know equations and from there I can estimate the projection matrix. So, this is one technique by which we can find out the projection matrix because we can always through experimentation, we can level some of the coordinate points in the world coordinate system and we can identify their image points in your in the images and that establishes those points correspondences then by knowing their coordinates, we can form this equations and then we can derive this elements of this projection matrix.

(Refer Slide Time: 08:58)



So, this is how these equations are written as I mentioned in the previous slide. So, you note here that a point in the image coordinate is given in a very general form as $(x_i \quad y_i \quad w_i)^T$ because it is denoting a column vector. You note here that we have used the scale factor w_i here. So, which means that this x, y is not exactly the observed image coordinate because there the scale factor has to be 1, but theoretically I can express an image coordinate in the using this scale factor in the projective space.

So, if I multiply an image coordinate point with this projection matrix, I will get a point in the 2 dimensional projective space and that is how in this form. So, the problem here is that if I apply simple equation then because of the scale factor that equation will be difficult to write t. So, we cannot simply equate the scales of PX_i and x_i. So, those are equated in terms of the after the scale adjustment those coordinates are their equivalence is established. Hence, $PX_i \equiv x_i$

So, instead of that we can consider them as vectors and these vectors so, their directions are to be same because the proportionate scaling does not change the direction of a vector. So, if I consider since they are equivalent in that case the interpretation is that those vectors they are parallel vectors they should have same directions. So, if I take the cross product of these two vectors and then form the equation because we know that the cross product of 2 parallel vector is a 0 vector.

So, we will perform this cross product and if we perform this cross product then we can derive these equations. So, let me find out that how this representation is possible. So, let us consider a representation of the projection matrix in this form which means I will be considering, now I will be considering the row vectors instead of column vector representations and because there are 3 rows.

$$P = \begin{bmatrix} r_1^T \\ r_2^T \\ r_3^T \end{bmatrix}$$

So, we have seen first row is represented by r_1, second row is represented by r_2 and third row is represented by r_3, I have used the transpose operations just to denote that in my actual presentations these are all column vectors, but since they are rows so, I have to apply transpose to those column vectors.

$$PX_i = \begin{bmatrix} r_1^T X_i \\ r_2^T X_i \\ r_3^T X_i \end{bmatrix}$$

So, now, if I perform PX_i so, the operation PX_i will give me the corresponding operation will give $r_1^T X_i$ then $r_2^T X_i$ and then $r_3^T X_i$. So, this is the column vector that we will get from this operation. So, you note that each one $r_1^T X_i$ what is the dimension of r_1, our dimension of r_1 is 4×1 because it is a row vector. So, each row is a 4 dimensional vector and what is the dimension of X_i, dimension of X_i is also 4×1. So, $r_1^T X_i$ will give you a scalar value, $r_2^T X_i$ will give you a scalar value and $r_3^T X_i$ give you a scalar value.

So, finally, you will get a in this form you will get a 3×1 column vector right. And X_i is represented in this way $(x_i \quad y_i \quad w_i)^T$ and if I take the cross product of this two as we did earlier in our first few lectures to get the cross product, we can write it as $r_1^T X_i$, this is $r_2^T X_i$ and this is $r_3^T X_i$ and then x i this is small x i y, i and w i. So, if I expand in this form. So, you will get the following vector

$$\begin{vmatrix} i & j & k \\ r_1^T X_i & r_2^T X_i & r_3^T X_i \\ x_i & y_i & w_i \end{vmatrix}$$
$$= (w_i r_2^T X_i - y_i r_3^T X_i)i + (x_i r_3^T X_i - w_i r_1^T X_i)j + (y_i r_1^T X_i - x_i r_2^T X_i)k$$

r 2 transpose X I, w I into r 2 transpose X i minus y I into r 3 transpose X i this is i plus let me write from this form.

So, X i into r 3 transpose X i minus w i into r 1 transpose X i that is j plus so, from surpassing this part. So, will be getting y i r 1 transpose X i minus x i r 2 transpose capital X i k. So, you will get this vector if I write this vector, we will get these vectors as w 1 r 2 transpose X i minus y i r 3 transpose. So, I have to use the other notation, this is not giving you the space, you can start again. So, w 1 r 2 transpose X i and then y i r 3 transpose X i so, this is the first component and other components and then that is equal to 0 0 0. So,

$$\begin{bmatrix} w_i r_2^T X_i - y_i r_3^T X_i \\ x_i r_3^T X_i - w_i r_1^T X_i \\ y_i r_1^T X_i - x_i r_2^T X_i \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

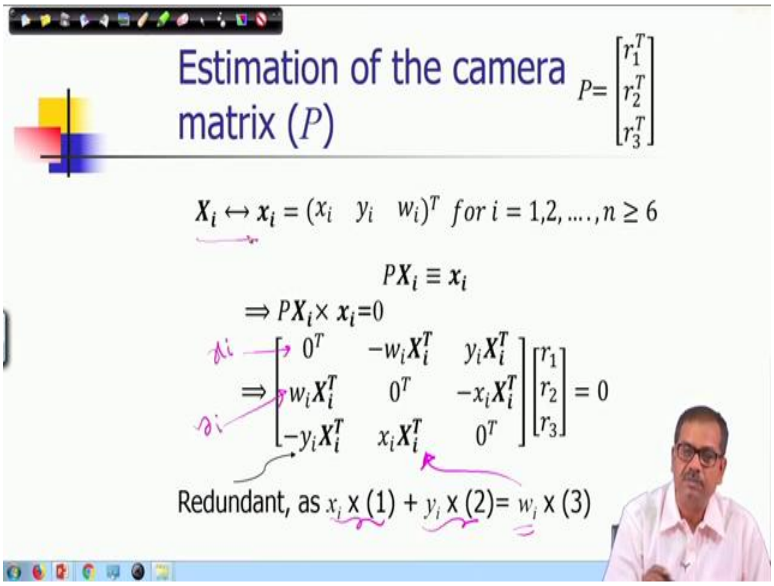
you will see that this is equated with this equation and here unknown is r_2 and r_3.

So, these are the unknown quantity because those are going to be estimated. Now, these equation is written in this form if you note the first row, you see it involves r 2 and r 3. So, it involves r 2 and r 3 since it does not involve r 1. So, this is multiplied by a 0 vector. So, you will see that 0 transpose here it is a 1 cross 3 0 vector. So, this is multiplied with r 1 which is 0 minus w i r 2 into X i transpose. So, now, since in my representation, I am representing the rows as a column vector, I will express all of them as a transpose operations.

$$\begin{bmatrix} 0^T & -w_i X_i^T & y_i X_i^T \\ w_i X_i^T & 0^T & -x_i X_i^T \\ -y_i X_i^T & x_i X_i^T & 0^T \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = 0$$

So, it is w 1 X i transpose. So, I will take it as a transpose operations so, it will be w 1 X i transpose r 2 minus y i X i transpose r 3 so, those equation. So, in this way the first equation is formed. Similarly, the second equation and third equation can be formed I have just shown only 1 equation just for the sake of your understanding. So, you can do it on your own and you can check how is equations are formed. So, now, you will see that there are actually three equations what you get, but out of them, there are two equations which are independent one is redundant, how it is so. So, let me show you.

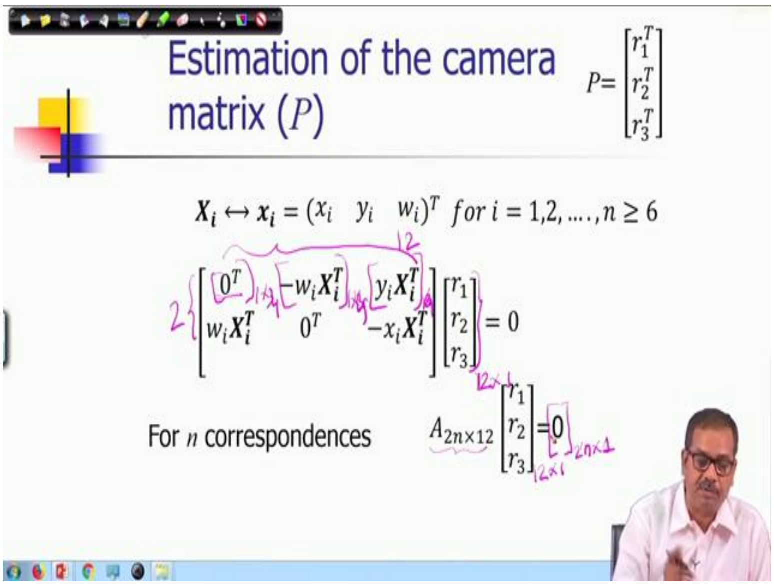
(Refer Slide Time: 17:38)



So, if I multiply the first equation with x i and then add it with and also multiply the second equation with y i. So, you multiply the first equation this equation with x i that is what is told here and these equation with y i and then if you add them you will get the third equation again which is multiplied by w i, you can check that thing. So, which means these equation can be derived from these two equations. So, which means that there are only two independent equations other equation can be derived. So, that is why in the previous case I mentioned that there are two independent equations that be that would be formed by using a one single point correspondences specification.

So, in this way you require six such point correspondences and you can form twelve equations independent equations and use it to derive the estimate this parameters of the projection matrix. So, let me continue this you know operation further.

(Refer Slide Time: 18:47)



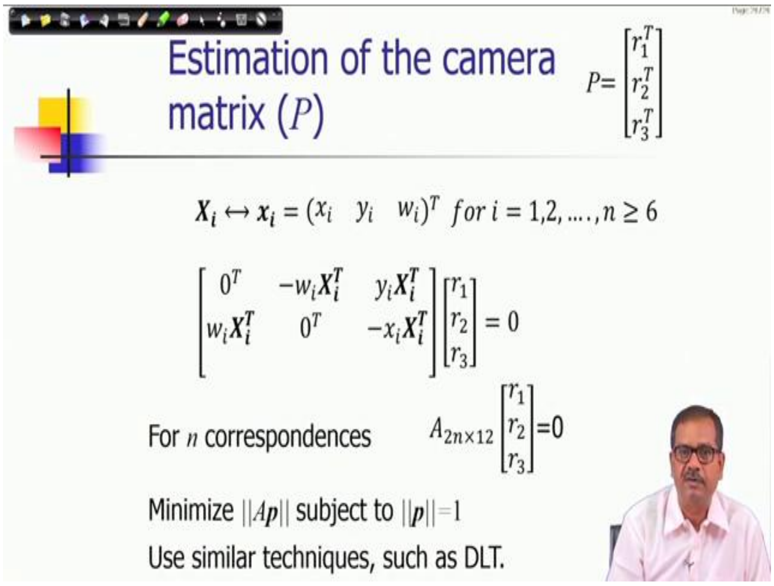
So, this is the summary finally, as you can see I have eliminated the third row just to show that there are only two equations which are formed by using a single point correspondence. So, if I have n such points, n point correspondences each one will give me two equations. So, I will get 2n such rows for n point correspondences and what is the dimension, if I note the dimension see each one is it is a 1×3

So, each one is a 1×3 sub matrix, each one is 1×4 sub matrix because number of elements here since it is a projection matrix, number of elements is 12×1 ok. So, you will get in this case, the total dimension here number of column is equal to 12 and this was 2 rows, for n point correspondences. So, you expect it should be $2n\times 12$ that should be the dimension of the matrix. So, you will get the equations in this form.

So, we will represent the corresponding matrix as A this matrix if I stack all this rows then we call this matrix as A which is of dimension $2n\times 12$ and each one each row is multiplied by this you know this elements of the projection matrix which are represented by their row vectors r 1, r 2, r 3 which is also of dimension $2n\times 1$ and each one each point correspondence will give you two equations.

So, finally, there would be this should be $2n\times 1$. So, this is a 0 column vector. So, this is a interpretation of this matrix representation of the linear equations in the matrix form. So, there are actually 2n number of linear equations when you get n point correspondences.

(Refer Slide Time: 20:59)



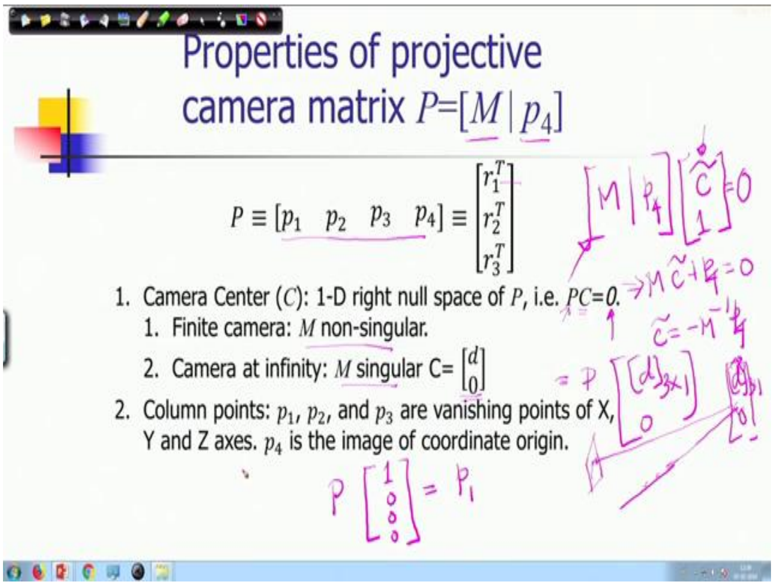
So, the as you understand that you have more number of equations than the number of unknowns so, you have to apply a least square error technique what we did earlier for estimation of homography also and in this case, it is a set of homogeneous equations these linear equations they form a set of homogenous linear equations.

So, the standard technique is that we need to minimize this objective function given that it is a norm of Ap, A is this corresponding matrix, A is the same matrix which is derived from the point correspondences and p is your solution; that means, the elements of the projection matrix which is represented in the form of this; that means, all row vectors are concatenated one by one from the starting from its first row to third row and that gives a 12 dimensional vector. So, we have to minimize this norm and since you know this solution is in the projective space and if I take a scaled vector also it is also the solution.

So, we would put a constraint on this minimization that norm of this vector should be equal to 1. So, this is once again the corresponding problem formulation for solving this particular problem for estimation of projective matrix and we have seen the same formulation for the estimation of homography matrix.

The same techniques we can use that is a direct linear transformation techniques by using either least square error estimation by considering one of the scale one of the; one of the element of the projection matrix of the vector is by setting it to some value and then we can solve it in the form of a non-homogenous equations by using least error technique or you can consider by finding out the eigenvectors of a transpose A and taking the vector corresponding to the minimum eigen value. So, we have already discussed in the previous class.

(Refer Slide Time: 23:14)



So, let me now consider that, once you have this projection projective camera matrix then what are the; what are the properties and what are the information that you can get by exploiting this properties so, we will be discussing those things. So, as we have discussed that projection matrix can be represented in different ways.

So, one of the representation is that it could be two sub matrices; one is of 3×3 M, the other one is a column vector 3 cross 1 and the notation we use for 3×3 sub matrix is M, the other one is p_4 or we can represent it as a set of four column vectors; that means, first column vectors, second, third, fourth that is the each one is a sub matrix of 3×1 or we can considered as a stack of rows where each row is of 4×1 sub matrix.

So, this is how we can represent a projection matrix with this notations and using this notations, we can relate these parameters to different types of imaging points so, different types of geometric points. So, first thing is the camera center, we have already established this relation . we know from the projection geometry that camera centre is a singular point; that means, if I would like to take the image of the camera center you will not be able to form a image, you cannot form a ray connecting to the same point itself.

So, which means that if I multiply the camera center with the projection matrix, I will get a singularity which is a 0 vector so that is the interpretation that PC should be equal to 0 and for finite camera, M is non-singular and for camera at infinity, you will find M is singular; that means, when camera is at infinity will again understand this interpretations.

So, where you can see that camera center is a point at infinity means its scale factor should be equal to 0. Now, if I consider this representation PC=0, I can derive the camera center very easily. So, using the sub matrix manipulation say I will represent the projection matrix in this form $[M \mid p_4]$ and the center as $\begin{bmatrix} \tilde{C} \\ 1 \end{bmatrix}$ where $\tilde{C}$ is the camera center in the world coordinate system and as I mentioned that should be equal to 0.

$$[M \mid p_4]\begin{bmatrix} \tilde{C} \\ 1 \end{bmatrix} = 0$$

Now, this if I perform the sub matrix you know operation multiplication operations, this is equal to $M\tilde{C} + p_4$ that is equals 0 and which means that $\tilde{C} = -M^{-1}p_4$ So, if I get the projection matrix elements, I can easily estimate the camera center by exploiting this relation where as for computing the center when M is singular then we have to use only M to find the 0 of that 0 vector of that M and to in the singular form and you can compute its center $C = \begin{bmatrix} d \\ 0 \end{bmatrix}$, we will again discusses this elaborate this computation later.

There are also interesting relationships with its column vectors so, you have the column vector say p_1, p_2, p_3 and they are vanishing points of X, Y and Z axes whereas, p_4 is the image of the coordinate origin. So, what is the vanishing point, suppose I consider particular direction or a particular ray.

So, the point which is at infinity that would be projected that would be also projected in the image plane. Now, it can be shown that for all the points lying in this straight line, they will be converging to that point and that is called vanishing point for that direction.

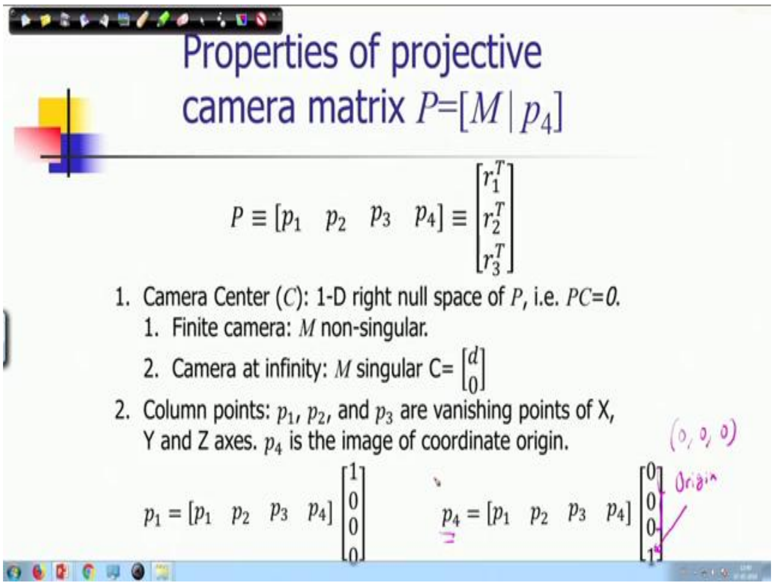
So, what you need to do, simply this representation this point representation would be the direction is given by the vector d.

So, we can represent this point as $\begin{bmatrix} d \\ 0 \end{bmatrix}$, you note that d is a vector, d is a vector it is a 3×1 element d is a vector. So, d is a 3×1 element and this 0 ok. So, if I multiply with respect to this projection if I multiply this vector with projection matrix, we will get the corresponding vanishing point of that direction. So, if I have the X-axis the direction of X-axis is given as $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and if I multiply p with this vector, what you will get, you will get only p_1.

$$P\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = p_1$$

So, p_1 is the vanishing point of the X -axis similarly p_2 is the vanishing point of Y-axis and p_3 is the vanishing point of Z-axis. So, this can be summarized very easily.

(Refer Slide Time: 28:56)



So, this is what I just described. So, p_1 if I multiply this with this, you will get p_1 and similarly p_2 p_3 and the other thing is that you note if I consider this point $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ and if I multiplied with the projection matrix, I will get p_4 which is the column vector.

$$p_1 = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

So, what is the interpretation here, what is this point? Note here this point this part is the origin of the world coordinate system because this is a 0 0 0coordinate system and this is the scale factor standard representation of the of any point in the homogenous coordinate space. So, $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$ this point is nothing, but the origin of the world coordinate system. So, if I take the image of the origin of the word coordinate system, I will get the column vector p_4 which means p_4 is the image of that coordinate origin. So, with this let me stop here for this lecture and then I will continue once again this topic in the next lecture.

Thank you very much.