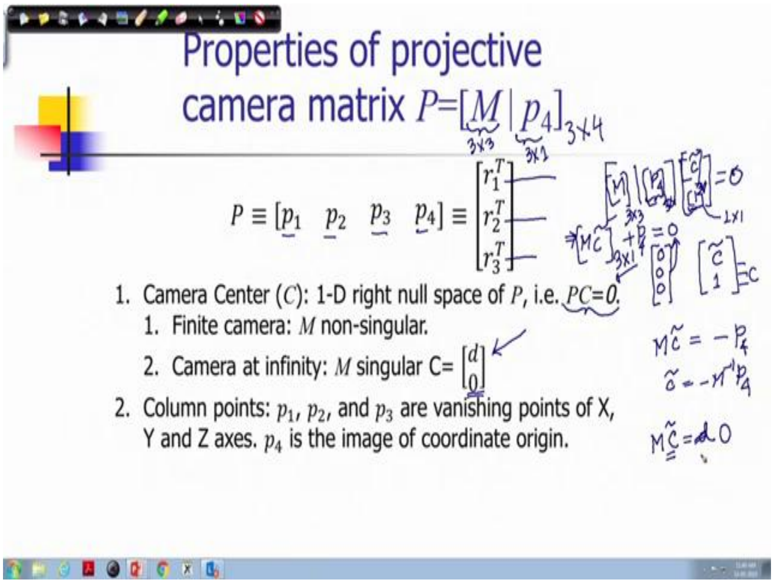
Computer Vision
Prof. Jayanta Mukhopadhyay
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

Lecture- 13
Camera Geometry Part- III

(Refer Slide Time: 00:20)



We will continue our discussion on single view Camera Geometry. And there in the last lecture we started discussing on retrieving various information from the projective camera matrix by exploiting its properties. So, we have already considered that how a projective matrix would be represented. So, there are various forms by which we represent a projective matrix. For the convenience of our discussion all these relations can be expressed in one of these forms.

So, the first from $P = [M \mid p_4]$ which you should note that a projective matrix which is a 3×4 matrix, it is represented by its sub matrix components. So, one of the representations in this form is that it has a 3×3 sub matrix, which is denoted here by the symbol capital M and there is a column vector which is the fourth column vector. So, we use this notation p_4, just to show that it is a column vector at the fourth position of the column vector, and it is a 3×1, its size is 3×1.

$$P \equiv [p_1 \quad p_2 \quad p_3 \quad p_4] \equiv \begin{bmatrix} r_1^T \\ r_2^T \\ r_3^T \end{bmatrix}$$

We can represent also projection matrices by their column vectors only. So, since there are 4 column vectors so, these representations like first, second, third, fourth column vectors; they are denoted by these symbols. And otherwise also we can represent them as a stack of rows, or where you have the first row and since our in your convention, we are representing a vector in the column form so we use transpose operation to denote that it is a row and so, r_1, r_2 anr_3d these are the three rows in this representation.

So, let us see how with this representations we can express various relations, with different parameters in the camera geometry or different information related to camera geometry; we will discuss those issues. the camera center that can be obtained as we also discussed these things in the previous lecture. The camera center has a property of becoming a right 0 vector of the projection matrix, because there is a singularity at camera center in the projection.

All the image points they are formed as an intersection of a ray which is connected to the center of camera and this intersection is with the image plane. But if the point itself is a camera center so, you cannot form a ray. So, that is the problem and that is why there is no such definition of image of a camera center and this expression mathematically can be translated in the following form. $PC = 0$

So, if I multiply the camera center coordinate in the homogeneous coordinate system, with the projection matrix P in this form then I should get a 0 vector.

So, you understand this 0 is not a single or scalar 0 value. It is a vector, becauseit is dimension should be $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ so, its(P) 3×4 and (C) 4×1 so, it is dimension should be $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ ok. So, that is a 0 vector. So, if I solve this problem how do you how can we express this camera center? I will consider this from suppose, the originals the center in the non-homogeneous coordinate system is denoted as $\widetilde{C}$, as we use this convention in our earlier lectures also. So, $\widetilde{C}$ and if I use a scaling coordinate as 1 in the additional dimension of homogeneous coordinate system so, this ($\begin{bmatrix} \tilde{C} \\ 1 \end{bmatrix}$) is representing the camera center.

So, I can write this relation as $[M \mid p_4]\begin{bmatrix} \tilde{C} \\ 1 \end{bmatrix} = 0$. So, which will be expressed if I perform the sub matrix multiplication with respect to this and you should note that when you are doing sub matrix multiplication, the first check should be dimensionality matching between the corresponding matrix multiplication. For example, M is a sub matrix of dimension 3×3 and $\tilde{C}$ is a column vector of dimension 3×1.

So, I can multiply M $\tilde{C}$ which will give me a column vector of 3×1. And then I will multiply p_4 and 1 so, it is the same matrix multiplication rules I am applying. So, I will be multiplying M with $\tilde{C}$ plus multiplying with p_4 with 1. So, p_4 as you understand p_4 is a its dimension is 3×1 and 1 is a scalar value its dimension is 1×1. So, I can simply write p_4, because see p_4 so it would be simply 3×1 matrix. So, it is a matrix multiplication is valid.

So, this should be equal to 0 which is also a 3×1 column vector. So, again I can reduce this equation. So, $M\tilde{C} = -p_4$ and then $\tilde{C} = -M^{-1}p_4$. So, you can see that you can get the camera center very easily by using this expression from the projection camera matrix. But note that M should be invertible then only you can get this expression. So, there are camera matrices where M is a singular matrix, which is not invertible. So, in that case camera center is not lying at any no physical; you cannot capture it in physical dimension, but you can represent it mathematically which is lying at a plane at infinity.

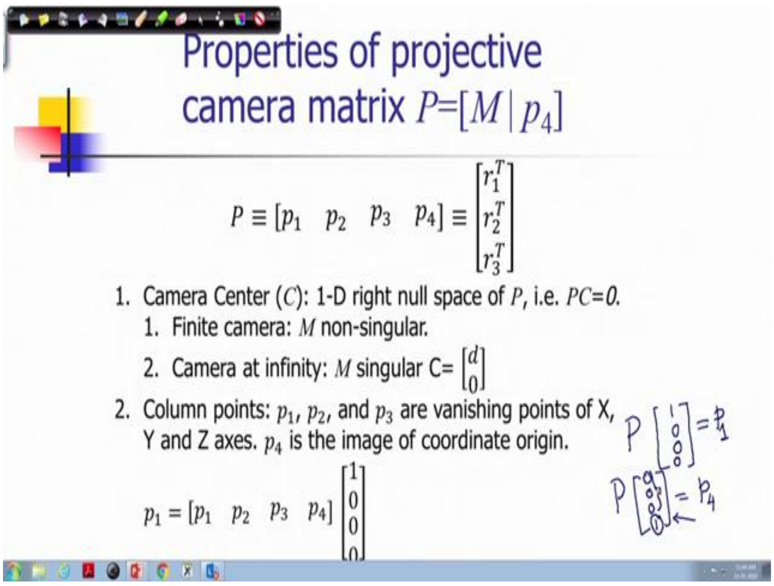
That is a concept we will discuss once again, but its representation any point in the plane at infinity is represented in this ($C = \begin{bmatrix} d \\ 0 \end{bmatrix}$) form. You see that it is a point once again in the homogeneous coordinate representation, this value of the scale dimension should be 0. So, that it is a point in the infinity. So, if I divide by 0 everything becomes infinity, but as I discussed earlier also the interpretation of any point t in this form which is at infinity is that it is a direction. So, it is just showing a direction. So, it is the point lying in that direction at infinity.

So, this is the form of the camera center in that case and this point can be computed once again by computing the right 0 of M, because M is singular there is a again a 0 vector. So, we can compute that value so, in that case we will have $M\tilde{C}$ should be equal to d. So, we can find out from the expression itself. So, when your center is in the form of $\begin{bmatrix} d \\ 0 \end{bmatrix}$.

So, this p_4 becomes 0. So, we will find $M\tilde{C} = 0$ and e by finding out the 0 vector we can obtain this 0 point. So, let me rub this value because this is wrong. So, it should be $M\tilde{C} = 0$ that we have to solve.

So, these are the things related to the camera center. Similarly, from the column vectors we can get some interesting informations. So, one of the information that we can get is the vanishing points of certain directions. As we discussed earlier also say, when we consider a point at infinity we should or you can choose you can choose any direction and a point at infinity can be represented in that directions by simply using the corresponding direction cosines as a vector and putting an additional informational dimension that homogeneous scaling dimension as 0.

(Refer Slide Time: 10:09)

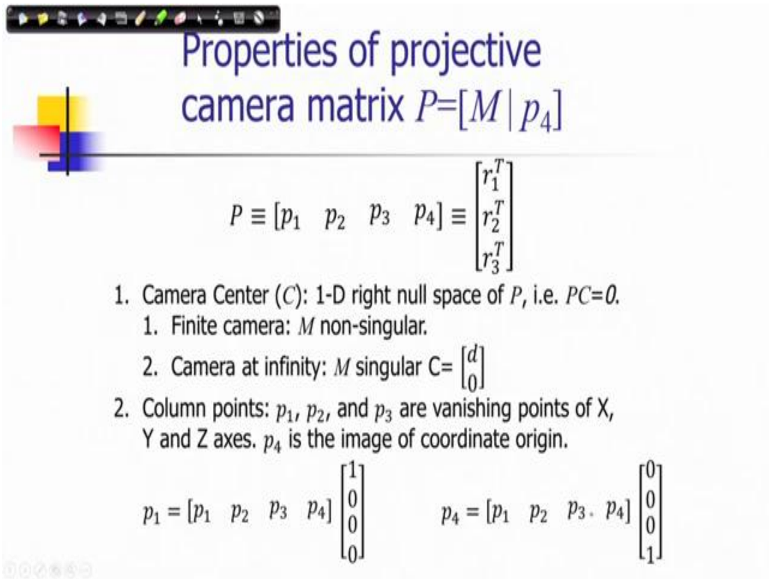


So, if I consider for example, X axis so, X axis direction is given by say $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and a point which is at infinity along that direction should be expressed by adding another 0, concatenating another 0 in the homogenous coordinate dimension ($\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$). So, if I would like to get the image of these particular point so, I should multiply with the projection matrix, then whatever value I would get in the homogenous coordinate system that is the image point to the corresponding point at infinity.

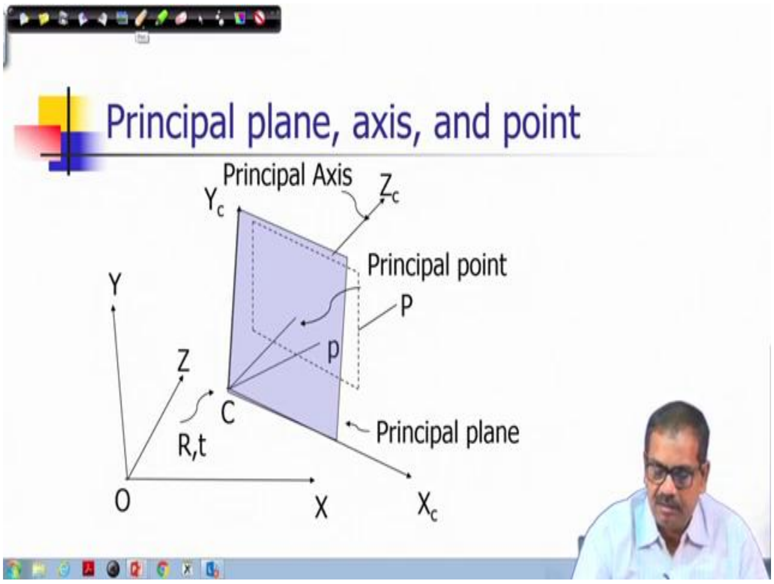
So, we call this kind of points as vanishing point. So, if you move towards that direction, all the images of points all lying on this particular (Refer Time: 11:02) line or will be converging to that vanishing point. So, in this case since P can be represented in the form of column vector so, n you can check that this operation will simply give me p_1. So, that is why p_1 is the vanishing point of X axis. Similarly, p_2 is also vanishing point of Y axis and p_3 is vanishing point of Z axes. So, this is one relation I just have shown here.

The other interesting fact here as you can see that what about p_4. So, p_4 you can get when you consider a point say $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$, if you multiply with the projection matrix P then you will get p_4. So, what is this point? This is not a point at infinity, because its scale value is 1 and this is a physical point in the space itself and what is this point in the three-dimensional space it is simply the origin. So, which means p_4 is a image coordinate of the origin. So, this is the other property so, this is about the p_4.

(Refer Slide Time: 12:16)



(Refer Slide Time: 12:20)



So, next we will be considering some of the; some of the geometric concepts related to this projection geometry, related to this imaging. We have defined all these you know concepts earlier also. They are principal plane, principal axis and principal point. So, in our notation as you can see that this is a very simple representation of the projection in a camera centric coordinate space where; p is a image of say a point in the scene represented by P.

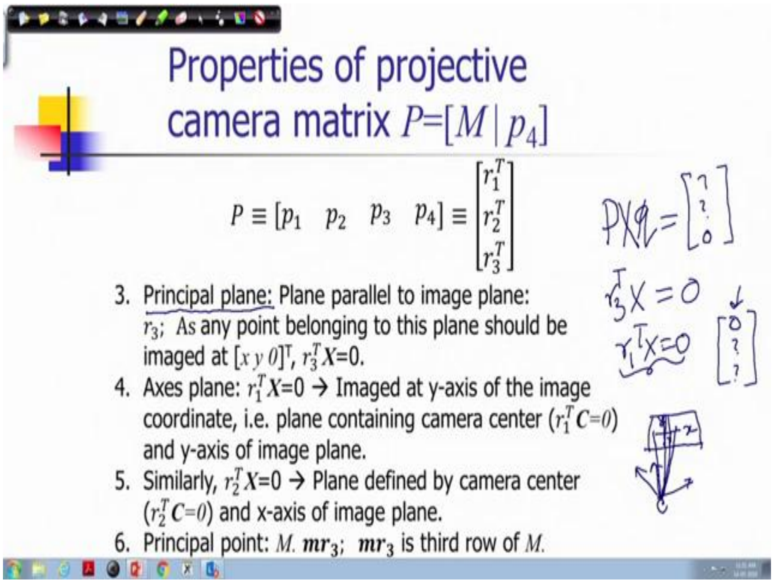
But no, externally you have another coordinate system, where we are observing the positions of these points and we call it world coordinate system. And the relationship between world coordinate and camera centric coordinate system can be established by applying simple rigid body transformation; where R denotes the rotation of axis which can be represented by a 3×3 rotational matrix and t denotes the translation of origin to the camera center, which is represented by a column vector or a just a 3 dimensional vector. So, this representation we have already discussed while deriving the projection matrix.

Now, we will be considering how to drive the relationships or information related to principal planes, principal axis and principal points. So, let me redefine once again. So, a principal point as you can see so, a principal points principal axis and principal plane; so, here principal plane is the plane where first let me define principal axis. So, principal axis is the Z axis in the camera centric coordinate system, which is incidentally the optical axis of the lens in the optical camera. So, we have considered that is a Z axis in our coordinate convention that is a canonical form.

So, in that way the Z axis is denoting the principal axis. So, intersection and principal plane is the XY plane of the camera centric coordinate system. So, XY plane is a plane parallel to image plane and that plane contains the center of the camera. So, geometrically that is the definition independent of any coordinate convention. So, it is the plane which is parallel to image plane and containing the center of projection. That is what is the principal plane. In this case incidentally in our canonical notations of coordinate system of camera centric coordinate system, principal plane is the XY plane. And principal point is the point of intersection of principal axis with the image plane.

So, these are the few definitions which we discussed earlier also and that I am providing here and now let us consider what are the informations that we can get. So, related to this principal plane others, once again we will be considering the representation of the projection matrix in this form and here I have shown those relationships.

(Refer Slide Time: 15:36)



So, one of the fact what is shown here that first how we can compute the principal plane. So, in this case you can see that since the principal plane is parallel to image plane. So, any point in the principal plane for any point you do not have any physical image point in that plane, in the image plane, because the ray forming from that point to the camera center does not intersect image plane. It intersects at infinity in that sense.

So, which means that the scaling dimension in the homogenous coordinate representation for the image point should be always equal to 0. So, if I consider any point say q which lies on a principal plane, if I multiply that point with these projection matrix, you will get an image point you do not know what these two things are there, but surely in the third place in the scaling dimension you will get it 0.

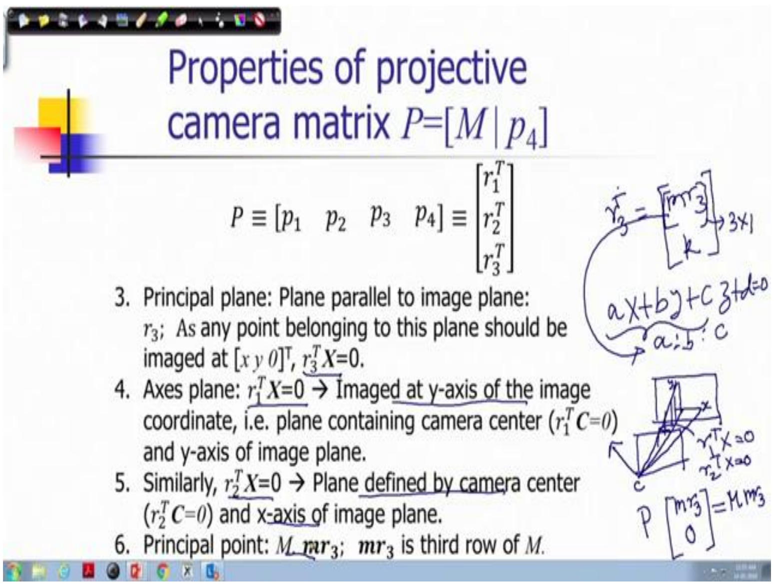
So, which means I can simply write in this case I have represented the point by X. So, let me use this symbol instead of q, let me use the point as X representation. So, if I consider multiplication with respect to the row vectors so, we will get $r_3^T X$ that should be equal to 0. Now you can see that this is nothing, but the equation of a plane and that is the principal plane.

You should note that this is what how do you should get a principal plane, but it is in the wall coordinate system that you should know. Similarly, if you consider this q there would be also image point where say for example; the first coordinate position would be

0 and rest other you do not care, which means if I multiply $r_1^T X$ that should be equal to 0 and that is a point that is a plane where all the points which are lying on that plane, they are projected in this form their image is in this form. So, what is this plane that would be interesting to understand, because in this case since your this coordinate which is X coordinate that is always equal to 0, which means you are considering only Y axis in the image plane.

So, if I consider your image coordinates convention and in image coordinate convention, you have X axis and you have Y axis whereas, in your camera centric coordinate system or if I say camera centric axis representation, say this is a center; so, if I consider this three-dimensional configuration and draw planes like this; let me rub it once again just to make it clear. Say you consider the image plane and this is Y axis and this is X axis and then you have a camera centric coordinate representation where C is represented. So, it is the plane formed by this axis Y axis and the camera center, in that plane whatever points are lying so they will be projected on Y axis. So, this is the where the X coordinate will be equal to 0.

(Refer Slide Time: 19:56)



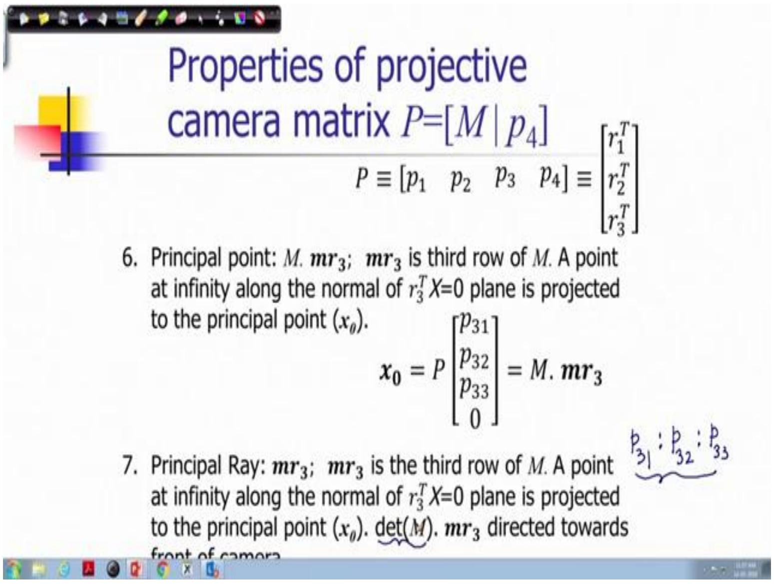
So, this is this plane is given by $r_1^T X = 0$. Similarly, if I can consider the X axis, this plane should be given by $r_2^T X = 0$. So, these planes are called axis plane as it is noted here, $r_1^T X = 0$ which is imaged at Y axis of the image coordinate and plane containing camera centered and Y axis similarly $r_2^T X = 0$ is plane defined by camera centered and X axis of the image.

What about the principal point? Now, the third row itself is showing you the principal plane. So, as we have mentioned this is a plane so, in the principal plane, so this is a principal plane; so this direction now r_3, if I consider the r_3 representation in this form it is a column vector r_3^T . So, let us consider that there is the it is represented as a 3×1 column vector and some scale value.

So, this mr_3 is a 3×1 column vector. So, as you understand from the equation of plane. So, if I write the equation of plane, in the form of say $ax+by+cz+d=0$. So, the normal of the direction of the normal of that plane is given by $a:b:c$ by this vector. So, mr_3 is actually given you giving you this direction.

So, what is the principal point? Principal point is the image of that point, which is lying in that direction and as I mentioned any point lying in at infinity in that direction should be represented in this form. So, in this case mr_3 is that direction and the point lying at infinity in that direction is 0 and then the image of that point is a principal point. It is nothing, but the intersection of the normal of the principal plane with the image plane. So, if I take the projection of this, if I take the image point so, this is represented as $M.mr_3$ That is what is shown here, then you will get the corresponding principal point. So, this is how this three important information you can get from the camera matrix.

(Refer Slide Time: 22:58)



We have discussed this principle point and here it is just explained once again, how this computations can be no it is elaborated with respect to this projection matrix. The mr_3 is further elaborated by the elements of the projection matrix in this form.

$$x_0 = P \begin{bmatrix} p_{31} \\ p_{32} \\ p_{33} \\ 0 \end{bmatrix} = M.mr_3$$

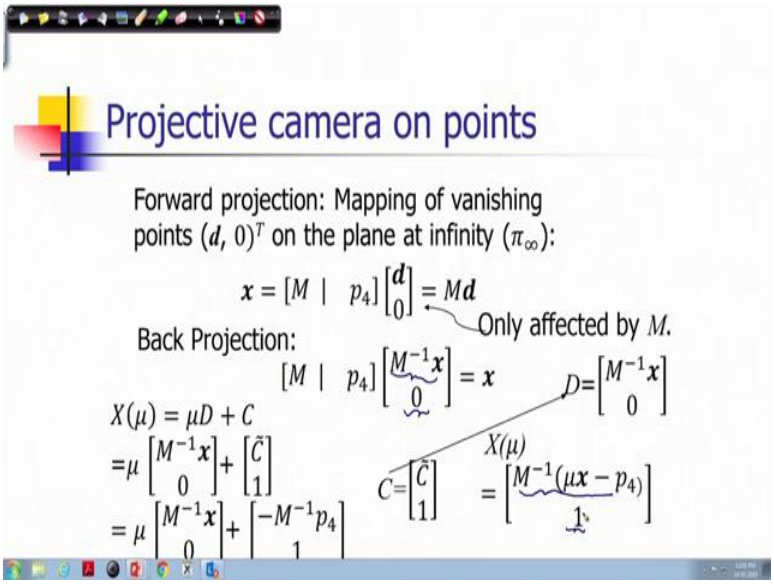
You note that p_{31}, p_{32}, p_{33} this these are the elements of the third row and that you need to know multiply with M, then you will get the corresponding principle point.

Similarly, the principle ray it is nothing, but the direction cosine given by mr_3 that is the direction vector, so, it should be proportional to p_{31}, p_{32}, p_{33}; I can consider it as

p_{31}: $p_{32} : p_{33}$. So, this vector, this direction is providing you the corresponding direction of the principal ray and now in the projective geometry in the projective sense, it could be in the along the camera centric coordinates Z axis or it could be in the opposite direction also.

So, from there we exploit a property of the transformation matrix. So, if you take the determinant of the transformation matrix and consider its sign. So, that sign will show the actual direction of this you know this ray. So, , this vector should be represented by the sign of the determinant and then the directions are given by mr_3. Instead of sign we can since it could be scaled, I can simply multiply the vector with determinant of the matrix M that is also possible.

(Refer Slide Time: 24:56)



So, let me then discuss another interesting fact that how the points which are there in the image plane, how they are back projected to form a ray. First let us understand the mapping of vanishing points on the plane at infinity. This we have discussed earlier also its respect to with respect to the vanishing points of principal axis like X axis Y axis and Z axis.

So, here if I consider any particular directions d, similarly we can multiply simply projection matrix with this vector $\begin{bmatrix} d \\ 0 \end{bmatrix}$ and that would give me;

$$x = [M \mid p_4]\begin{bmatrix} d \\ 0 \end{bmatrix} = Md$$

that would give simply multiplication of the matrix M with d. So, that is the vanishing point of the directions are the points at infinity lying in that direction d. So, you should note that vanishing point is only affected by M.

So, the back projection can be defined as the ray formed by the camera center with the image plane and you can express that ray in the world coordinate system. That is how the back projection is expressed. So, if I consider a point say $\begin{bmatrix} M^{-1}x \\ 0 \end{bmatrix}$, which means it is a point which is lying in the direction of $M^{-1}x$.

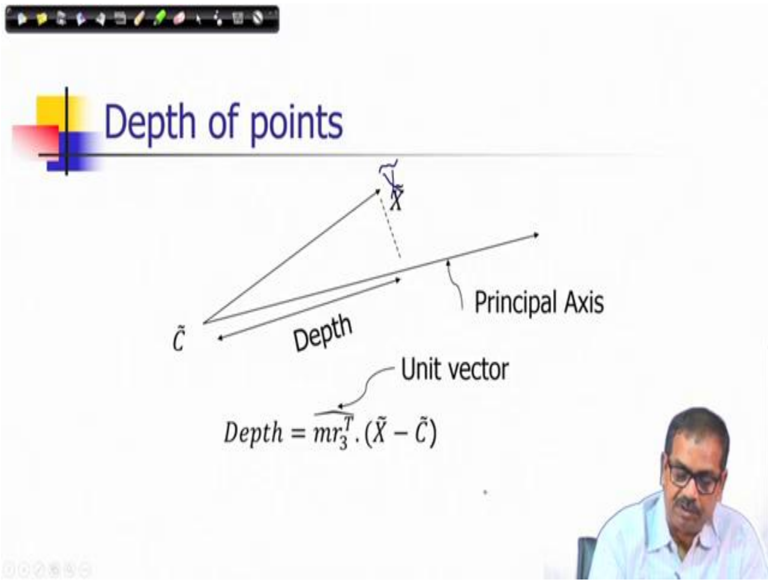
So, this is the directional component in that direction and it is a point which is lying at infinity. This 0 is showing then the image of that point is x which means this is actually the ray, projection ray, the direction of the projection ray, it is expressed in that particular fact. So, finally, we can form a ray which passes through a center, say $\widetilde{C}$ in the homogeneous coordinate system is to represented in this form and at infinity the point is expressed as $\begin{bmatrix} M^{-1}x \\ 0 \end{bmatrix}$, because the direction is $M^{-1}x$. So, this can be easily expressed by a parametric form of equation of straight line in a three-dimensional space where, $M^{-1}x$ is the direction and a point on that line is defined by $\widetilde{C}$.

So, this is how the parametric form is expressed as you can see, that it is a linear combination of the point in the homogeneous coordinate system. The point at infinity in that direction D in that direction of $M^{-1}x$ and $\widetilde{C}$ and following is the expanded from.

$$
\begin{aligned}
X(\mu) &= \mu D + C \\
&= \mu \begin{bmatrix} M^{-1}x \\ 0 \end{bmatrix} + \begin{bmatrix} \widetilde{C} \\ 1 \end{bmatrix} \\
&= \mu \begin{bmatrix} M^{-1}x \\ 0 \end{bmatrix} + \begin{bmatrix} M^{-1}p_4 \\ 1 \end{bmatrix}
\end{aligned}
$$

In fact, this ($\begin{bmatrix} M^{-1}(\mu x - p_4) \\ 1 \end{bmatrix}$) is the form which is more convenient to express this $\begin{bmatrix} M^{-1}(\mu x - p_4) \\ 1 \end{bmatrix}$. So, these are the points which are lying on the projection ray connecting to the cameras center.

(Refer Slide Time: 28:00)

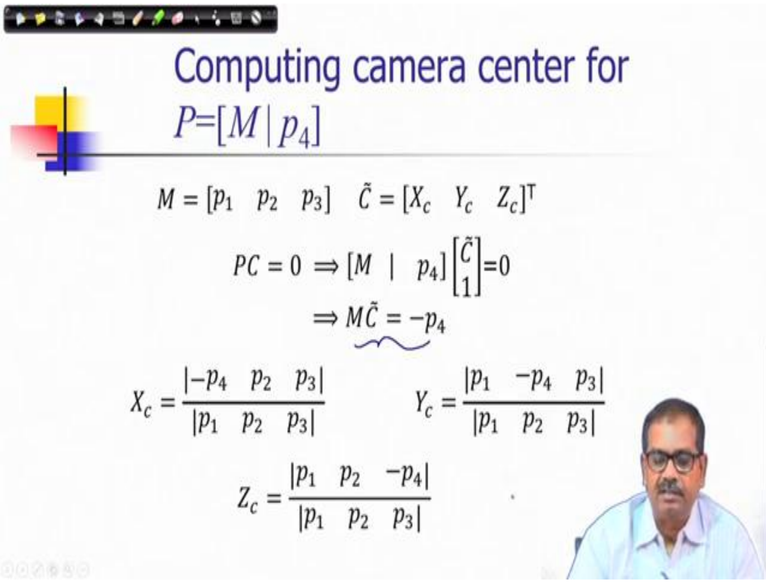


How do you compute the depth of points? So, once again depth of point is nothing, but the projection of a projection ray connecting to the scene point and on the principal axis. So, you consider any particular scene point say X, $\tilde{X}$ here in the three-dimensional space. So, its projection so, this vector its projection on the principal axis which is Z axis of the canonical form a rod in the camera centric system. So, this distance is the Z disk. This distance is the depth information. So, this is a depth so, this could be easily expressed by this form.

$$Depth = mr_3^T.(\tilde{X} - \tilde{C})$$

So, mr_3 is the corresponding direction. So, tilde on top of mr_3 denotes that it is a unit vector along that direction. So, if you take the projection means you have to take the dot product along the unit vector of that direction. So, this vector is defined by $\tilde{X} - \tilde{C}$. then you get the corresponding depth information. So, this is what its unit vector.

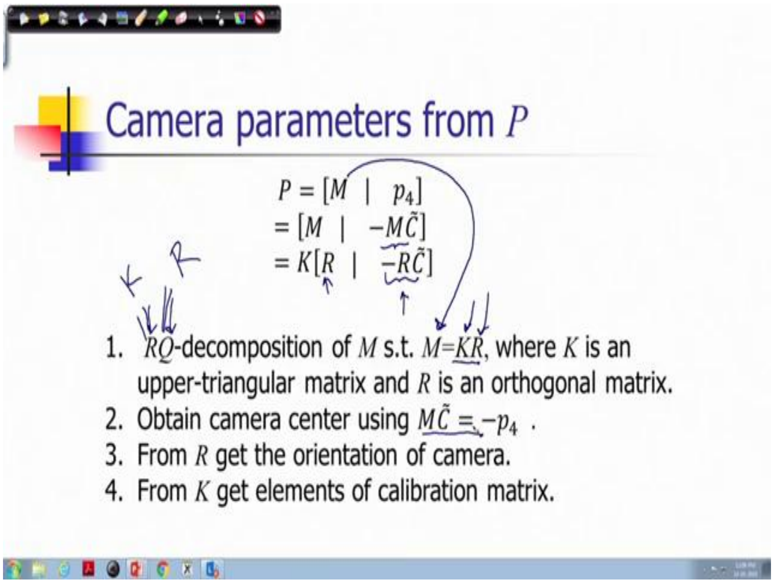
(Refer Slide Time: 29:13)



There is another easier representation of camera center by solving the set of equations expressed in this form. As we discussed earlier also, you can see that in fact, there are three equations, because camera center is represented by three coordinates. So, you need to find out three coordinates. So, this could be easily expressed by using crammers' rule and the representation of M in terms of column vectors.

So, let me show you the solution. So, X_c that is a X coordinate of the camera center is given in this($\frac{|-p_4 \quad p_2 \quad p_3|}{|p_1 \quad p_2 \quad p_3|}$)form by using the Cramer's rule. So, this is a determinant formed by the matrices of the column vectors, determinant of the column vectors similarly Y_c is given by this ($\frac{|p_1 \quad -p_4 \quad p_3|}{|p_1 \quad p_2 \quad p_3|}$) and Z_c is given by this ($\frac{|p_1 \quad p_2 \quad -p_4|}{|p_1 \quad p_2 \quad p_3|}$). So, you apply simply Cramer's rule and then you can get this expression.

(Refer Slide Time: 30:15)



So, next we will be discussing how you can get the camera parameters from the projection matrix P. So, here in this case we are again representing projection matrix in different form, like you have a $P = [M \mid p_4]$ or $\left[M \mid -M\widetilde{C}\right]$, which($-M^{-1}\widetilde{C}$) is also representing p_4 or you consider a decomposed form of M, where you have the camera calibration matrix and the rotation matrix that is $KR = M$ then, this part is $-R\widetilde{C}$. All these representations have been discussed in previous lectures.

So, one of the easiest way to get the camera parameters is by using matrix decomposition, because you know the M in general this representation is formed by the multiplication of these two matrixes K into R. So, K is the camera calibration matrix and R is the rotation matrix of the transformation from world coordinate system to camera coordinate system and the property of K it is an upper triangular matrix and R is an orthogonal matrix.

So, if I apply RQ decomposition where R is a upper triangular matrix and Q is an orthogonal matrix, if I perform this decomposition, then I will get two matrices which satisfies this property. Now this solution may not be unique in that sense, but still we considered that this is one possible solutions through matrix decomposition.

So, we consider after decomposition, this R is equated with K and it is an upper triangular matrix and this Q is equated with R; note the you know notational little ambiguity here, but in this case this is a standard term of decomposition. So, this is the form we are using. So, from there you can get the parameters of calibration matrix, from the upper triangular matrices. Similarly, from the rotation matrices you can get the parameter of rotations and by operating camera center as we discussed earlier, you can get the corresponding translation parameters also.

So, this is how you get all the parameters of the camera matrix. So, here we will stop our lecture at this point and we will continue this discussion in the next lecture.

Thank you very much for your listening.