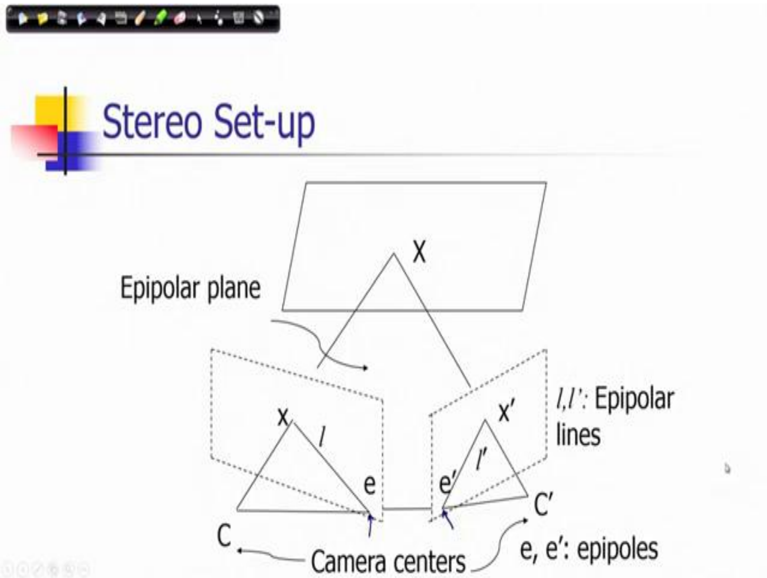
Computer Vision
Prof. Jayanta Mukhopadhyay
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

Lecture - 16
Stereo Geometry Part – 1

In this lecture, we will start a new topic that is on Stereo Geometry. So, let us consider what is meant by stereo setup. In the stereo setup, we have two cameras and in this two cameras, they are specified by their camera centers and their image plane.

(Refer Slide Time: 00:37)



And scene point X for which you will be taking the images in both this cameras. So, if I apply the rule of projection here. So, we get the image of this scene point X in the first camera, whose camera center is given as C as a say x.

Similarly the second camera given by the camera center C' forms an image of that same scene point at in its own image plane. So, you have two images of the same scene point and this is what y x' ou will get in a stereo setup. For every scene point, there are two image points, now if they are viewable from the two cameras itself.

Now, this particular construct has certain properties geometric properties. What we can see particularly that the point C , X , C' and the if I consider a plane found by C C' and the scene point X . So, the image points x and x' , they also lie on the same plane.

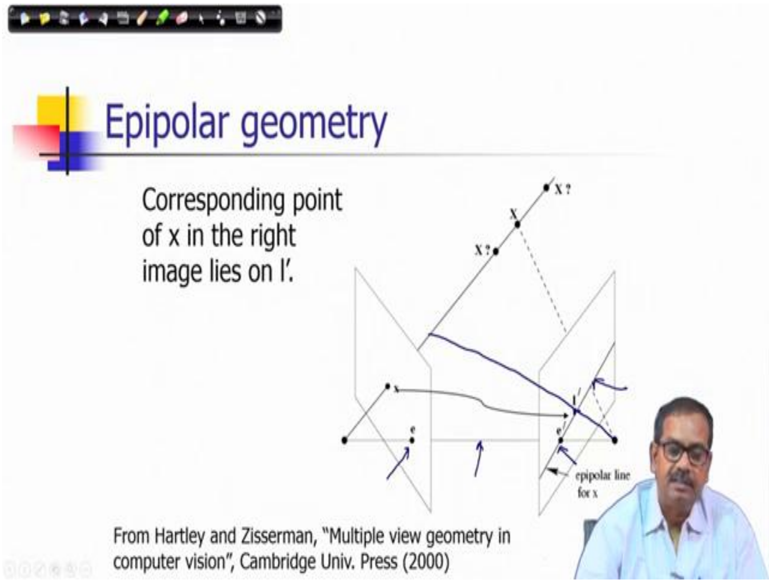
There is a from first restriction, also if I draw a line between C and C' which is called for a stereo setup as the base line. So, that also lies on the same plane; that line intersects the image planes..

Again, there are some image points in the image plane and in fact, as we can see that one of these intersection point is here which can be interpreted as the image of camera center of the other camera. And the other point is shown here point of intersection of the base line with the other image plane, which we can interpret also as image of camera center of the first camera.

So, these two points, these are called as epipoles and they are also lying on the same points. So, they are denoted here by e and e' and the line joining e and x which is nothing but intersection of the plane containing all these points X , C , C' and their image points. So, that intersection with the first image plane, it will be giving you another line and similarly get two such lines in two image planes because the intersections with respect to the three-dimensional plane formed by same point and camera centers.

Now, those two point; those two lines they are called epipolar lines. So, we have this particular configuration and these terms are used we will be explaining these terms more and more in subsequent lectures. But you should note this interesting configuration that all of this points and lines, they lie on the same plane and that plane is called epipolar plane with respect to this same point. So, you have this epipolar lines and this plane is called epipolar plane, which is formed by the camera centers both the camera centers and a same point.

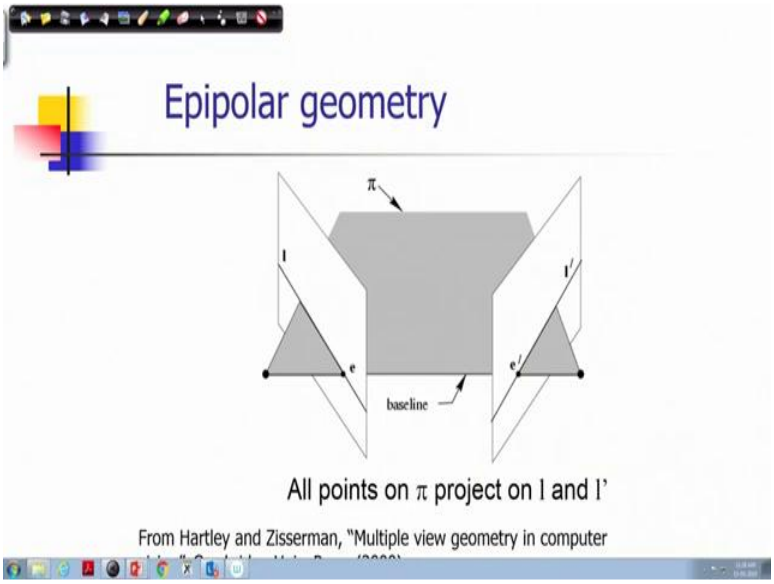
(Refer Slide Time: 04:09)



So, this geometry in general is referred as epipolar geometry. So, that summary of this again; once again different concepts of the epipolar geometry is shown here. As you can see that you have the corresponding base line here and the intersection of base line is giving in the epipoles and you note particularly this epipolar line in the second camera. So, if I consider the projection ray formed by the first camera and the image point, then any point lying on that ray is a possible candidate of a same point from which whose image is this x and if I take the image of the scene point by the second camera.

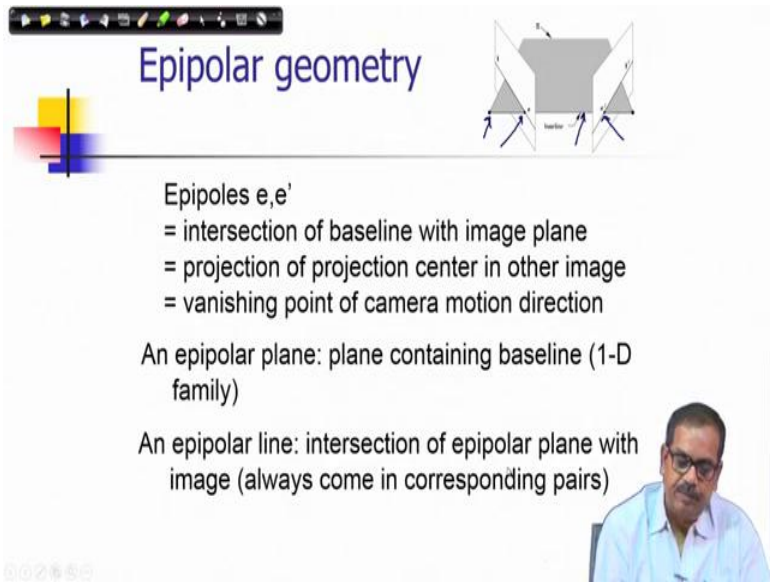
So, this image would lie on the epipolar line. So, that is a geometric interpretation. who all the possible candidates of same point with respect of this image point, they lie on an epipolar line. So, this is a constraint that is imposed by this particular configuration. So, corresponding point of X in the right image lies on l prime.

(Refer Slide Time: 05:29)



And you can also on interpret this configuration in this fashion also for any epipolar plane π, its point of intersections are those two; its line of intersections with respect to image plane. Those are the corresponding epipolar line. So, all points on π the project on those lines only. So, any point on plane π will be projected for image for the first camera will be projected on l; epipolar line l and for the second camera, it will be projected on epipolar line l'.

(Refer Slide Time: 06:09)



So, in an epipolar geometry we have this concepts once again we will be now defining them in particular using the geometry configurations only. So, epipoles as we see e and e' in the figure here. So, epipoles are defined as the intersection of baseline with image plane that is one kind of definition. So, as you can see that this is a baseline, this is a baseline and intersection with image planes respectively, they provide you those points of they are the epipoles. That is one kind of interpretation or you can also interpret as I mentioned earlier, that projection of projection center in other image because you can consider for example, this is the camera center which is a projection center of the first camera.

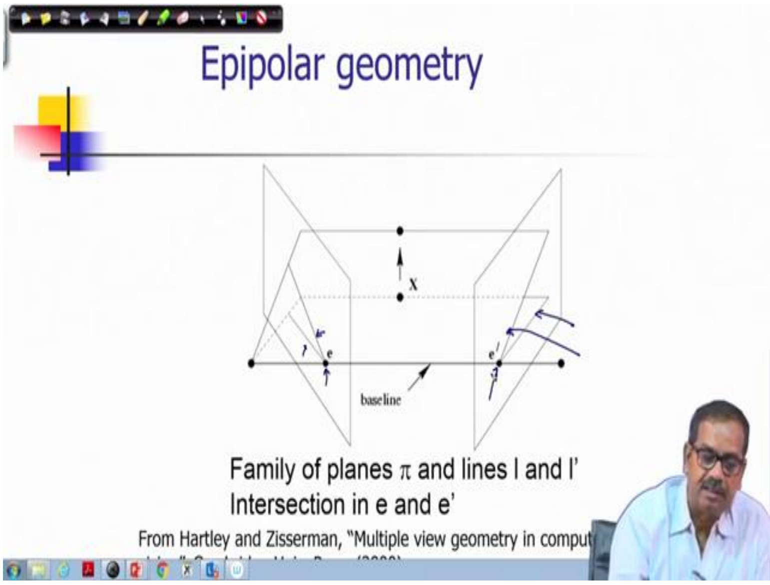
So, these epipole is the projection of this camera on this image plane because you draw the ray from this point to this point which is again the baseline and its point of intersection of the image plane will give you these particular epipole. We call this epipole as a right epipole. So, in our convention we will have this particular directionality, we will be using this directional notations. So, first camera we will call left camera from the diagram's representation . So, you will also sometimes refer it as left camera and the second camera, we will refer it as right camera.

Similarly, the epipoles out of this first epipole e is considered the left epipole, that is just the epipole on the reference image plane or first image plan. So, first camera is considered as a reference camera of the stereo setup and second epipole, it sometimes it is also referred as right epipole that is on the corresponding epipole on the second camera. So, these are certain notations also or terminologies will be used in our discussion.

You can also interpret epipoles as a vanishing point of camera motion direction. So, what is a camera motion direction? It is a direction of translation of the camera center. So, which is again given by the baseline. So, in that direction as you can see that it is the intersection of that ray with respecting image plane will give you that intercepting point will be always the vanishing point along that direction. So, epipoles are also vanishing point of camera motion direction. So, in various ways you can interpret this epipoles.

On the other hand, you define an epipolar plane it is the plane containing base line. So, you can see that baseline. So, it is an one dimensional family in the sense that it is a pencil its baseline is acting like a pencil and you know all the planes over centering around that particular line is defining an epipole; epipolar plane. And epipolar line is a intersection of epipolar plane with image and it always come in corresponding pairs.

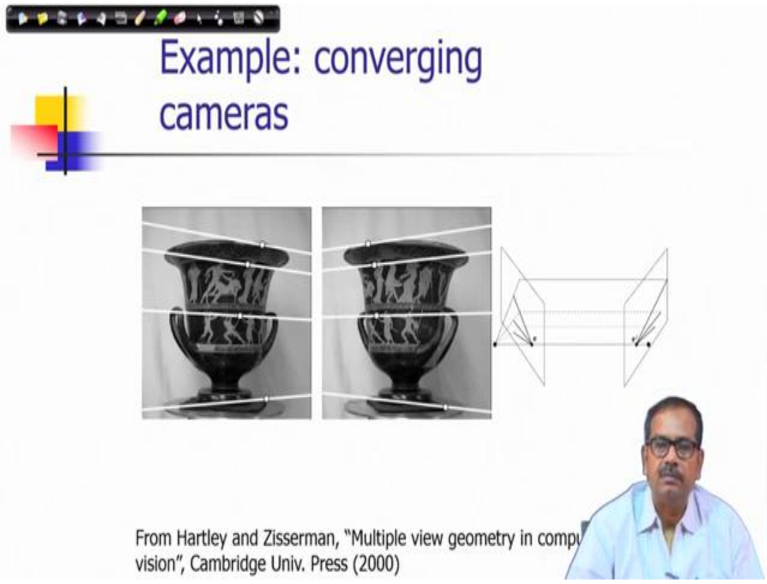
(Refer Slide Time: 09:25)



So, this is a explanation that was mentioning that you get a family of epipolar planes, intersection of each any pair of the plane is a baseline. And this epipolar planes that intersects with the image plane to provide you epipolar lines and all those epipolar lines, they are also meeting at epipoles in respective epipoles that we can see in this particular diagram.

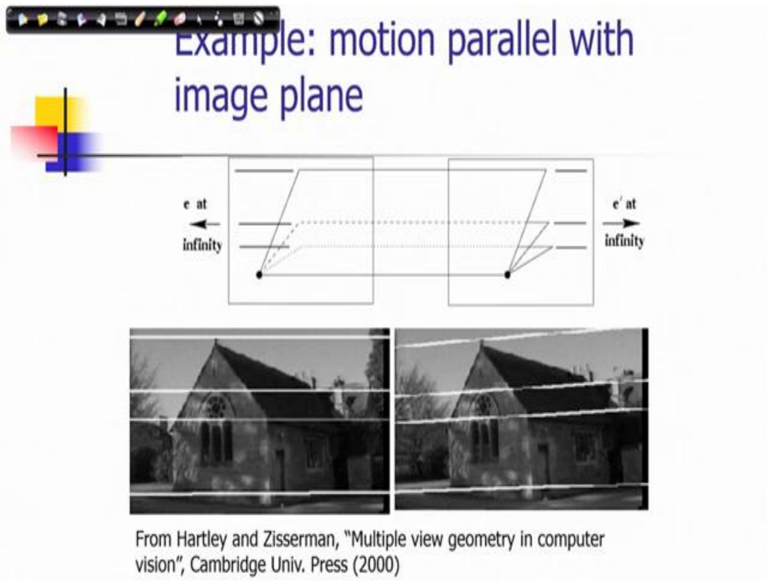
So, you can consider. So, this is a plane and this is a another plane. And for each plane, now these are the corresponding epipolar lines and they are intersecting at e in this case were intersecting at e'.

(Refer Slide Time: 10:12)



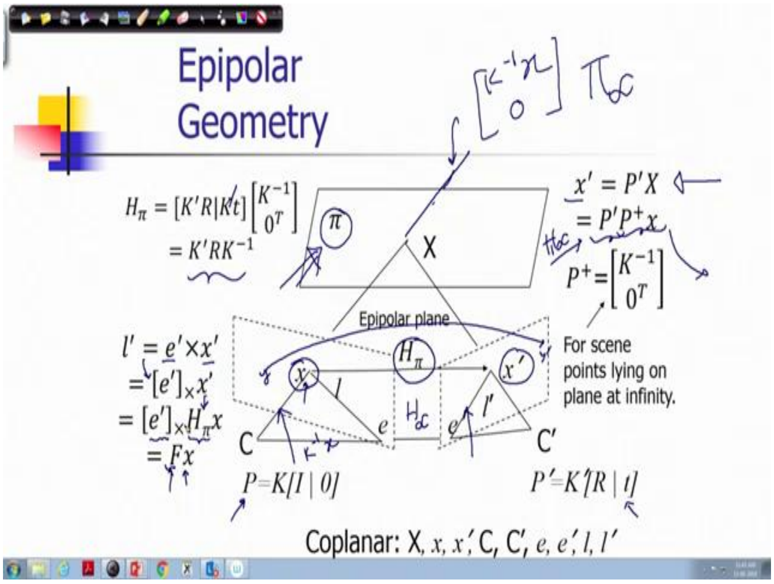
So, typical examples of converging camera images. So, were it is shown that their epipolar lines are converging. So, this is one example, I have taken from the book of Hartley and Zisserman which is also shown here multiple view geometry in computer vision. You can find, you can see that how epipolar lines they can be mapped on those images and you can find out those all those epipolar lines, they are converging in this case.

(Refer Slide Time: 10:46)



Also you take a situation, when you have simply translation of camera center. There is no rotation of axis. So, which means your epipolar lines they become parallel so, all epipolar lines they become parallel; parallel to the direction of motion of the image planes or the camera. So, since epipolar parallel lines, they meet at infinity. So, this epipoles also in the image plane, they are the points which are lying at line at infinity. So, these are the interpretations of epipoles when you have simply a translation of camera centers.

(Refer Slide Time: 11:30)



So, now let me provide you the mathematical formulation of different concepts involving these different entities which we have defined like, first thing as we can see from this geometry. Once again, that you have two cameras given by the centers C and C' and also the projection matrices of the each camera, those are also specified here. So, one of the projection matrix that is given as $K[I \mid 0]$; so that means, it is a reference camera and it is a camera centric coordinate system that we considered here. And on the other hand, the other camera that is the second camera is given by this ($K'[R \mid t]$) K' and then, the rotation and translation form as R and t.

So, the in a inwards studio set up, the reference camera is denoted in this diagram in this form. So, one of them is a reference camera and the other one is a second camera. This is a notation or left camera usually we put the reference camera call it as a left camera and the right camera as the other camera with respect to the reference camera. Its coordinates are expressed and reference camera is usually they have this camera centric coordinate system, but it could be also generalized.

Now, you can see that in this structure, we have two image points corresponding to the same points that is x and x'. So, we can actually show that given all the points in this plane, it induces a homography. We will discuss it later on, but right now we will consider a particular type of homography, we will consider the mathematical relationships and show that how that homography exists between corresponding points.

So, you know this particular relations up camera geometry that x' can be expressed as $P'X$ in the homogeneous coordinate system in the projective geometry. And then, we can consider the ray found by this particular image point with respect to camera center and consider a point which is lying at infinity along that ray. So, that is expressed as pseudo inverse P plus x ($P^{+}x$).

So, I should not call exactly that is pseudo inverse, but it is given in this structure $\begin{bmatrix} K^{-1} \\ 0^{T} \end{bmatrix}$.

You can note that the direction cosine, direction of the particular line, this can be given as $K^{-1}x$ and the any point which is lying at infinity for this line that can be expressed as $\begin{bmatrix} K^{-1}x \\ 0 \end{bmatrix}$ ok. So, finally, this is a same point which we were considering which is at a infinity and we are considering an image of that point in this particular configuration.

So, if I consider for every image point for any other image point say y, we have also there is a corresponding image point y' which is again formed by the scene point which is lying at that infinity and that plane is called plane at infinity. So, there is a particular notation for that and all these points. So, this plane let us assume in this case that we will only considering the formation of corresponding points with respect to plane at infinity, then we can get an interesting homography between those corresponding points.

That means, , the point in the first camera first image plane and the corresponding points which are found from plane at infinity from the same points of the plane at infinity the second camera and this particular homography, we also sometimes referred as homography at infinity. So, this discussion also we will be doing later on, but let this structure of at infinity simple because as you can see from this relation x' and this is nothing but it is giving you the corresponding H_α and which is easily computed in this form ($P'P^+$).

So, this is what scene points lying at plane at infinity. So, I can simplify this matrix by its corresponding elements. So, this P' is given by $K'R$ and it should be $K't$. So, if I multiply, then you get the homography matrix as $K'RK^{-1}$. So, this is your point x', now the these line is reconstructed as I know in the projective space, it should be cross product of e and x' that would give you this line $l'=e'\times x'$.

Now this operation these are all three vectors and we know how to compute this cross product of this operation also can be expressed in terms of matrix multiplication. I form a three dimensional matrix with the elements of e' and I can express this particular cross product as I mentioned that with respect to multiplication of that matrix. So, let me provide you a simple notation.

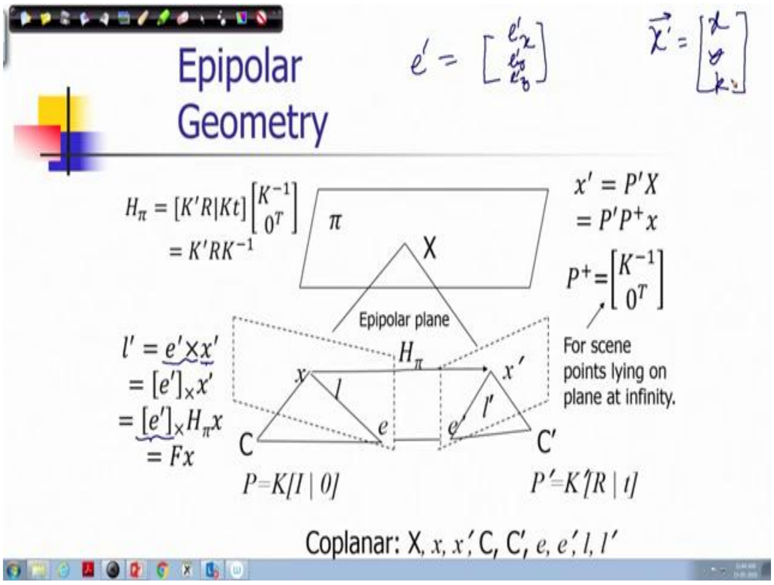
So, just to complete this discussion, we will come back to these particular constraints. Suppose there is a matrix let us say 3×3 matrix and it is multiplied with x prime that is giving you l prime which is performing equivalent computation of this cross product of 2 points. So, I can then expand this operation. So, this x' can be expressed as also the using the homography relation in this form; $H_\pi x$ and since this is a 3×3 matrix, this a 3×3 matrix, again I can multiply this two matrices in composite form and let me denote it by F ($[e']_\times H_\pi$).

So, this relation l' this epipolar line in the second camera can be obtain from the image point in the first camera. It is a very important relationship in this stereo setup; is a very fundamental relationship in a stereo setup and that is they are related by this particular

Expression ($l'=Fx$). So, there exist a matrix which can be obtained from this you know from this parameters, from this values and if I multiply the image point in the first camera in its homogenous coordinate system, with this matrix F which is of dimension 3×3, then I will get this epipolar line l'.

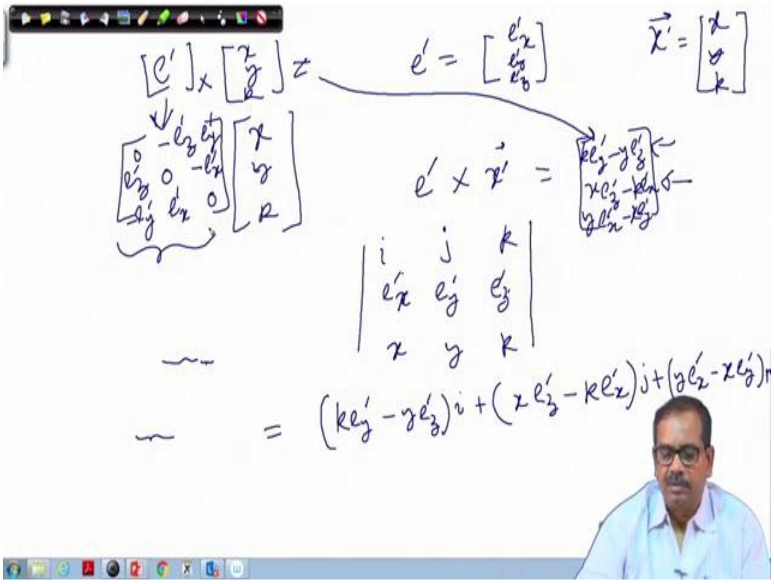
So, this matrix is called fundamental matrix and that is a very characterizing property or characterizing matrix of this stereo setup and its task is that it converts an image point to a line which is an epipolar line and sometimes they are called also, this relationship is also called co-relation. This transformation is called co-relation. That is a kind of misnomer, but you may find in the text book this term is also there.

(Refer Slide Time: 19:58)



So, let me also explain how this particular conversion means happening here, that a cross product can be expressed as the form of a three-dimensional matrix multiplication. So, if I consider e'. So, let us consider e' is represented by the column vector say $\begin{bmatrix} e_x \\ e_y \\ e_z \end{bmatrix}$ say let us consider these are the column vectors. And a point x' is: as you know that this is a basically a column vector ($\begin{bmatrix} x \\ y \\ k \end{bmatrix}$) and is its coordinate, let me represent it as say x y and some scale value k that is the homogenous coordinate system. So, let me represent in this form.

(Refer Slide Time: 21:05)



So, when we have a matrix multiplication; so, let us consider this scenario. So, when you have this , not the cross product of $e'\times x'$. So what we will do?

$$\begin{bmatrix} i & j & k \\ e'_x & e'_y & e'_z \\ x & y & k \end{bmatrix}$$

$$(ke'_y - ye'_z)i + (xe'_z - ke'_x)j + (ye'_x - xe'_y)k$$

We will consider this i j k and this is e x prime e y prime e z prime and this is say x y k. So, I need to compute this expression. So, let me expand this expression as we did earlier also. So, we can write it as say k e y prime minus y e z prime, this is i plus then x e z prime minus k e x prime, this is j plus y e x prime minus x e y prime, this is k. So, you get this particular expression. So, you would like to have a vector which give you all these things; that means, it should give you this is

$$\begin{bmatrix} ke'_y - ye'_z \\ xe'_z - ke'_x \\ ye'_x - xe'_y \end{bmatrix}$$

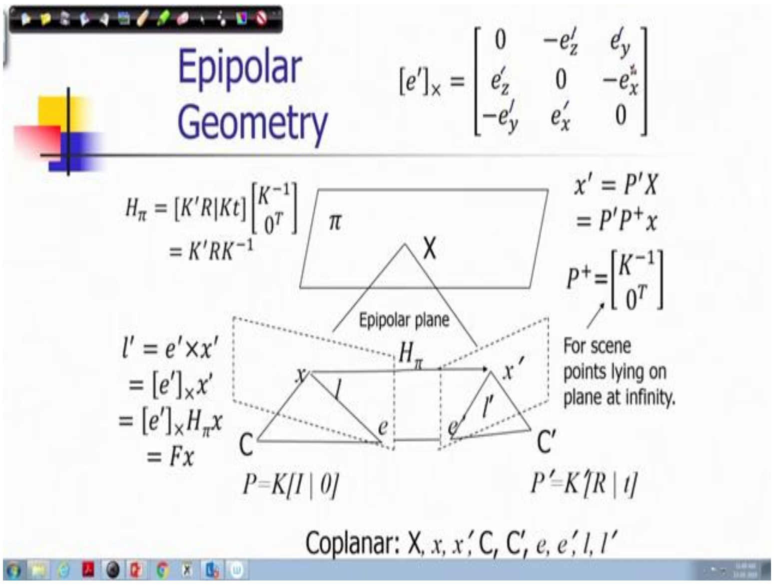
k e y prime minus y e z prime, then x e z prime minus k e x prime and y e x prime minus x e y prime.

So, I should have a matrix with respect to in a matrix multiplication, as I mentioned that we are considering replacing the as a matrix multiplications. So, it should multiply with $\begin{bmatrix} x \\ y \\ k \end{bmatrix}$ which should give me this particular expression. So, I have to find out this value.

Now, you can see that to get this what I have to do? So, there is no x component in this row. So, this should be 0, then for y it is $-e'_z$; for k it is e'_y. Consider the second row so, the x contains e'_z, then this is 0 and then with k, it is $-e'_x$ and consider the third row, where x contain $-e'_y$, then y e'_x and this is 0. $\begin{bmatrix} 0 & -e'_z & e'_y \\ e'_z & 0 & -e'_x \\ -e'_y & e'_x & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ k \end{bmatrix}$

So, this is what is a you know conversion of $e' \times x'$, I can as well replace it by the three-dimensional multiplication. You note particularly these matrix is a skew symmetric matrix. Its diagonal all the diagonal elements are 0 and you can find out the respective transpositions will give you a negation of the particular matrix elements. So, this is how this is explained in this case.

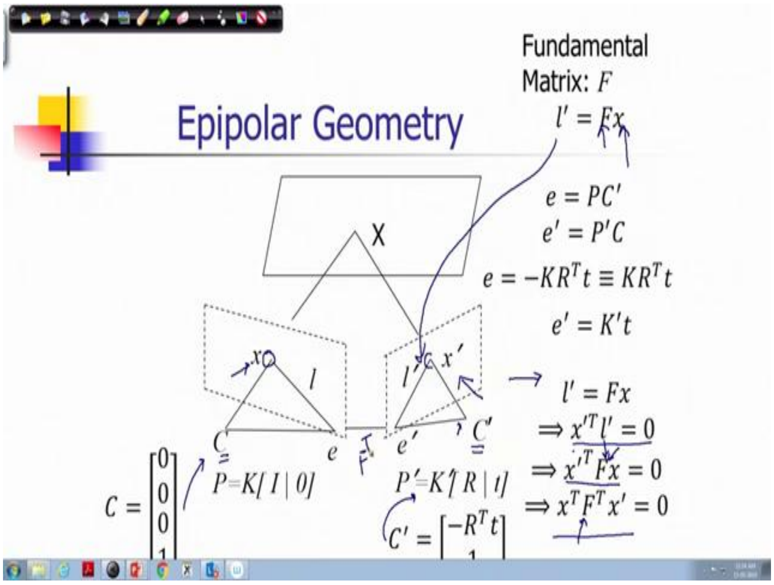
(Refer Slide Time: 25:05)



So, we will proceed with this particular explanation that how it could be converted and as I mentioned earlier that how this relationships, matrix relationships can be expressed into form of a skew symmetric matrix. Now you can use prime(‘) in my notation, I have used prime(‘) every case because they are all related to e'. However, you know it is just a matter of notation that you need to be used to.

So, fundamental matrix as you can see the expanded form, when I am considering only the elements of projection matrices and the epipoles that can be written as $[e']_{\times} K'RK^{-1}$. So, I can express this fundamental matrix by the camera matrices and epipolar configurations, from the epipoles itself.

(Refer Slide Time: 25:56)



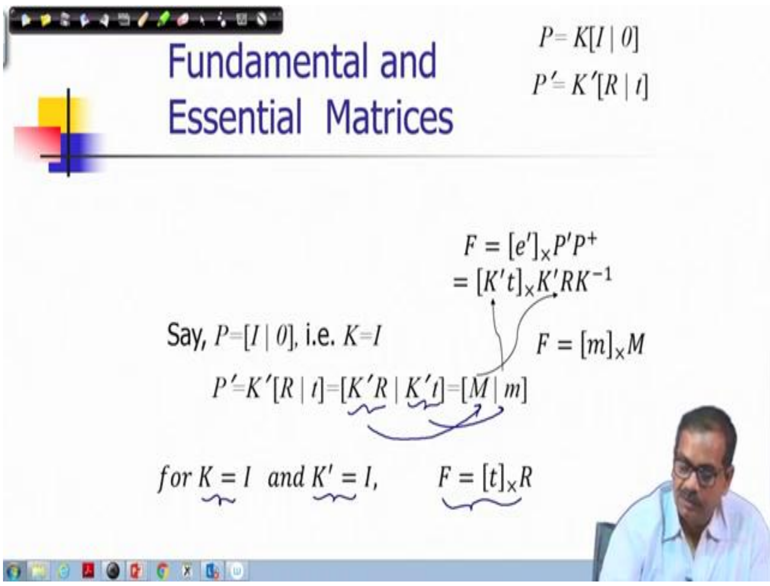
So, we will continue this discussion on epipolar geometry, once again that fundamental matrix is the matrix, if I multiply any image point in the projective space in the first camera, if I multiply that with this matrix F. Then, I will get the epipolar line in the second camera. So, the interpretation is that the image point so, all my corresponding image point any or rather corresponding image point will lie on this epipolar line. , which means that I can apply the point contentment relationship with line, you note that this center of the camera is given by this, center of the first camera is the given by this one and center of the second camera is also given by this particular expression. So, this is $\begin{bmatrix} -R^T t \\ 1 \end{bmatrix}$ so, given by this expression.

And as we have discussed that epipole can we considered as image of the camera centers. So, for the left epipole e, it can be expressed as if I multiply the projection matrix with the C', I will get the image coordinate of e and which is given by $-KR^T t$. So, which is $KR^T t$ or this will equivalent because you know directionality does not matter in this case; minus sign just denotes the directionality. Similarly, e' can be considered as the image of the image of the first camera center e' which is $K't$. So, e is image of C', e is image of C' and e' is image of C and this is how the relationship is there and the point contentment relationship with the epipolar line that is expressed here.

So, I can get the epipolar line l' that can be expressed as Fx and if I apply the point contentment relationship $x'^T l' = 0$ and expand l' in the form of equal to Fx. So, we get a relationship $x'^T Fx = 0$. So, given two corresponding points which are observable or measureable or computable with two images; that means, they are the points of the same point. Then, their existed matrix 3×3 matrix which satisfies this relation.

So, this relationship is also a very fundamental to epipolar geometry and infact. Now if I apply the matrix transposition rule, I can also express this is as $x' F^T x^T = 0$ that means, if I consider this is a reference camera and this is a second camera. So, with respect to that point correspondence also we have a matrix which satisfy that relationship and those matrices are called fundamental matrix and they are related by this relation.

(Refer Slide Time: 29:38)



So, if this is a reference camera with respect to this camera, if the fundamental matrix is F, if I consider the other camera is reference camera, then the fundamental matrix will be

F^T . That is what we get in this relationship. There is another particular term which is used, terminology which is used in epipolar geometry when the camera matrix when the camera is a calibrated camera.

So, we say a camera is calibrated, when the calibration matrix is known to us and then, we see that how this relationship can be simplified. So, when you have a calibration; when you know the calibration matrix, I can express any image coordinate into its normalized canonical coordinate system.

So, you can see that the relationship between the fundamental matrix and the camera parameters is given in this form ($F = [e']_\times P' P^+ = [K't]_\times K' R K^{-1}$). Given the two cameras as P and P' as shown here and suppose P is given as simply in a canonical form ([I|0]), normalized canonical form the calibration matrix itself ; that means, we have a apply all necessary transformation to bring it into the standard pinhole camera geometry; the base pinhole camera geometry configurations with the base coordinate convention of cameras centric coordinate system.

So, if I apply once again the representation of P' in this form, then this relation can be expressed as so, you can see that K'^T can be also considered as this is a representation.

$$P' = K'[R \mid t] = [K'R \mid K't] = [M \mid m]$$

This is an another representation of camera system and if I put the calibration matrices, I since if I know the calibration matrices I can always covert them by doing the necessary transformation.

Then, this particular structure will relate to t. There is a translation parameters $[t]_\times R$. So, this is how the fundamental matrices expressed in that form and then, this fundamental matrix is called essential matrix. In this particular form, this fundamental matrix is called is essential matrix.

So, with this, I will stop my lecture, this lecture here and we will continue this discussion.

So, thank you for listening to this lecture.