Computer Vision and Image Processing – Fundamentals and Applications

Professor. Doctor M. K. Bhuyan
Department of Electronics and Electrical Engineering
Indian Institute of Technology, Guwahati

Lecture No. 05

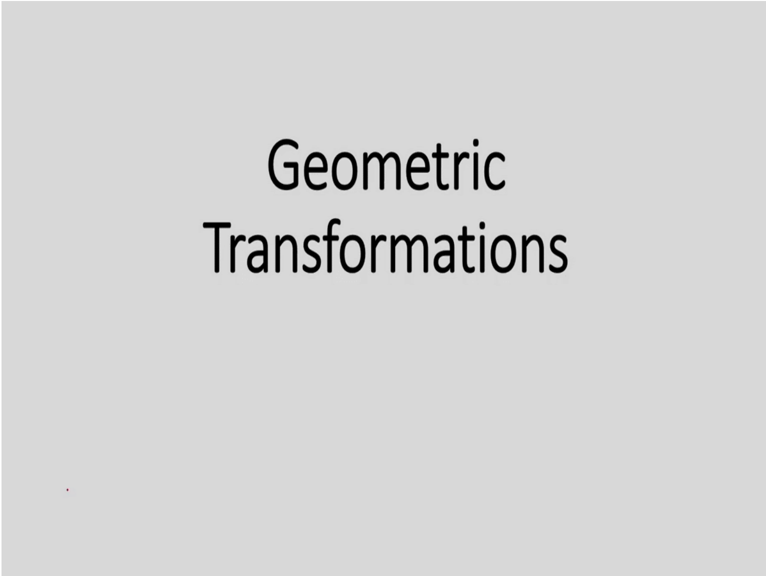
Image Formation: Geometric Camera Models

So, welcome to NPTEL MOOC course on Computer Vision and Image Processing Fundamentals and Applications. So, in my last class I discussed about the concept of shape from shading. So, image acquisition process is nothing but the 3D to 2D transformation. So, in this process the depth information I am going to lose. So, depth information is not available. So, that is why we considered the concept of shape from shading. So, in shape from shading I want to get the shape information from the shading information.

Shading means the variable levels of the darkness. So, one algorithm for solution of the shape from shading problem is photometric stereo. In photometric stereo we consider number of images. So, from number of images we want to get the shape information that is very important, the photometric stereo. So, suppose if I want to determine the shape information of a surface that surface is illuminated by number of light sources, but one source at a particular time. So, in this process I am getting number of images. So, from these images I want to get up shape information. So, that is the objective of the photometric stereo.

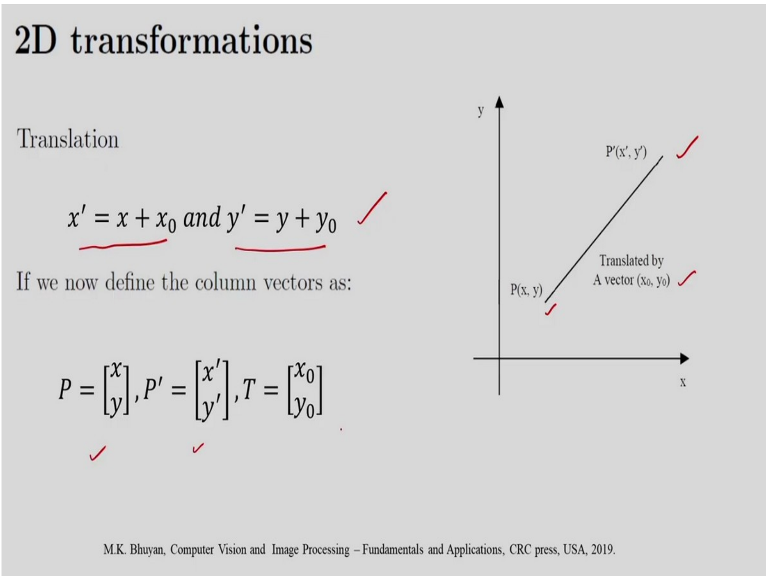
So, today I am going to discuss about the image formation concept. In this case I will discuss the concept of geometric camera models. So in this case, for understanding of the geometric camera models I have to discuss the concept of geometric transformations, because the concept like these perspective projections, orthographic projections, if I want understand, the concept of the geometric transformation is quite important. So, let us see what is geometric transformation?

(Refer Slide Time: 02:11)



So, suppose if I consider one object, suppose the object is moving in an image. That means it is translated or it can be rotated or it can be scaled. So, in this case these operations I can represent mathematically. So, if I consider combined operation, suppose translation, rotation and the scaling, so this combined operation is called the affine transformation. So, how to represent these operations mathematically? So, today I am going to discuss about the geometric transformation. And after this I will discuss the geometric camera models. So, let us discuss the geometry transformations.

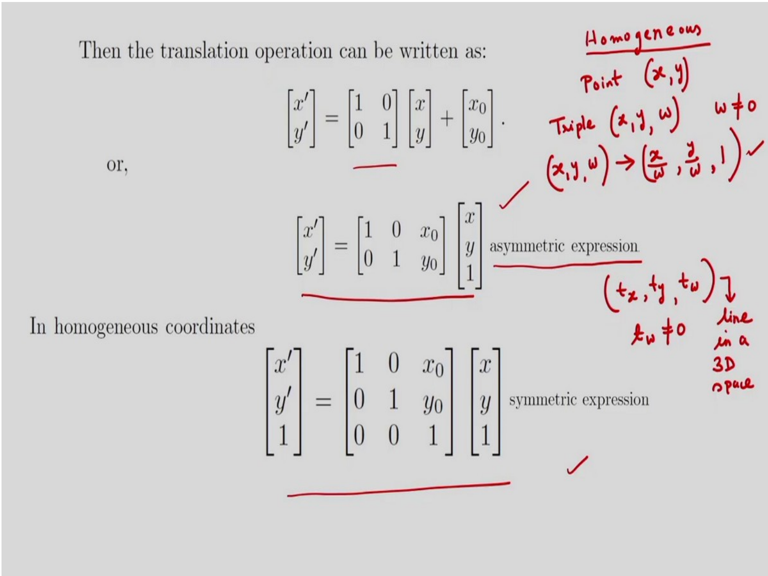
(Refer Slide Time: 02:49)



So, first operation is 2D transformations. So, first I will discuss 2D transformation and after this I will discuss 3D transformations. So, first I am considering one point here the point, the point is P, x y. This point P, x y is translated by a vector, the vector I am considering, that is x_0 y_0. So, it is translated by the amount x_0, x_0 unit parallel to x axis, y_0 units parallel to the y-axis.

The new point will be P' ,x' y' so, this is the new point. This translation operation I can write in this form. So,x' = x + x_0 and y' = y + y_0. So, this is a translation of a point. Suppose if I define the column vectors like this. So, P is the point, the initial point P x y. P' is the final point, x' y'. And in this case if I consider, the translation vector is T, x_0 y_0.

(Refer Slide Time: 3:49)

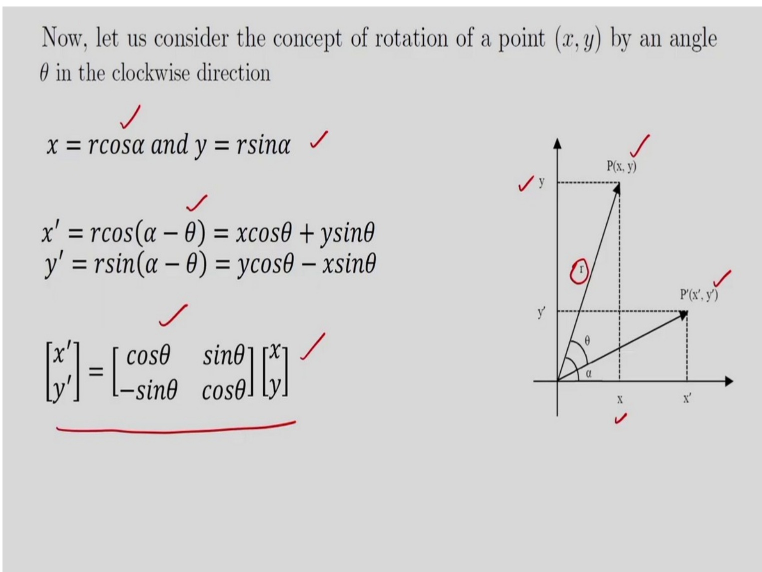


Now translation can be represented like this. So, translation process is x' y'= 1 0 0 1 hat matrix I am considering, into x y + x_0 y_0. So, that the previous equation you can write in this form. Or you can write in this form. x' y' is equal to 1 0 0 1 x_0 y_0 x y 1. So, the previous equation, whatever I have shown in the previous slide, that the translation operation, that operation you can show in the matrix form like this.

But one thing is this expression is asymmetric expression. So, there is a concept of homogenous coordinate system. So, the concept is homogenous, homogenous, homogenous coordinate system. Suppose if I consider a point, point is represented by the pair of numbers x y. So, each point can be represented by triple. What is a triple? x y w, suppose. If suppose w coordinate is non-zero then the point can be represented as, the point is mainly the x y and w, that point can be represented by x / w , y / w and 1.

These x /w, y / w, these are mainly the Cartesian coordinates of the homogenous point that means triples of coordinates represents point in a 3D space. The concept is something like this. I am considering Tx, the point, the triple Tx Ty and Tw.Tw is not equal to 0. So, in this case all triples of the form, this form actually represents a line in the 3D space. And if I consider homogenous coordinate system then in this case it is the symmetric expression. So, I can use symmetric expressions in place of asymmetric expression.

(Refer Slide Time: 06:07)

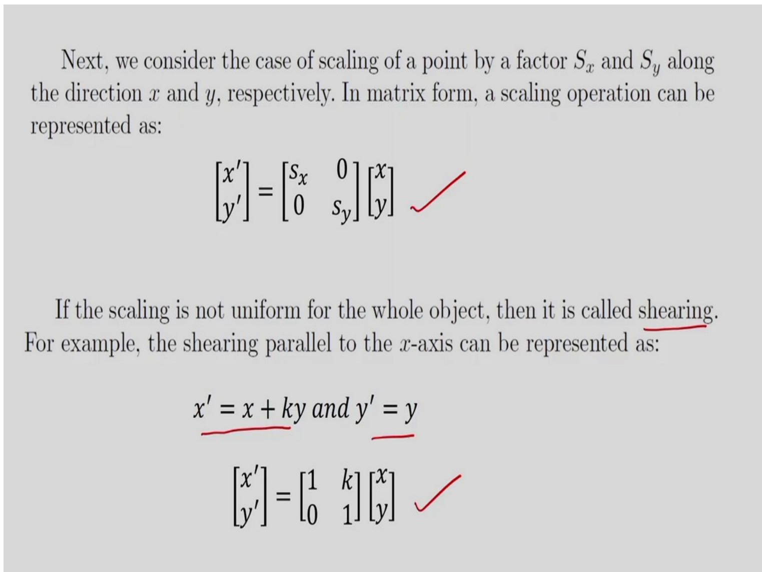


Now let us consider concept of rotation of a point x y by an angle θ in the clockwise direction. In this figure I have shown. Suppose the point is P x y and that point is rotated by an angle θ. So, what is the new position? The new position is $P' x' \; y'$. That point P x y is rotated by an angle θ in the clockwise direction.

Now the simple, the trigonometry, if you see what will be the x? What will be this x? x is nothing but r cos α because the length of this vector is r. So, the, it will be r cos α. And what will be the y? This is y. So, y will be equal to r sin α. So, from this you can determine x'. This x' coordinate you can determine. That is r cos α - θ is equal to, you can just expand this one, x cosθ + y sin θ and y' is equal to r sin α - θ = y cos θ- x sin θ.

Then in this case this rotation operation you can represent in this form, x' y' is the final coordinate. x y is the original coordinate, initial coordinates. And I have the transformation matrix. This is the transformation matrix, cos θ sin θ - sin θ cos θ. So, this matrix is the transformation matrix for rotation. The rotation operation you can represent it like this.

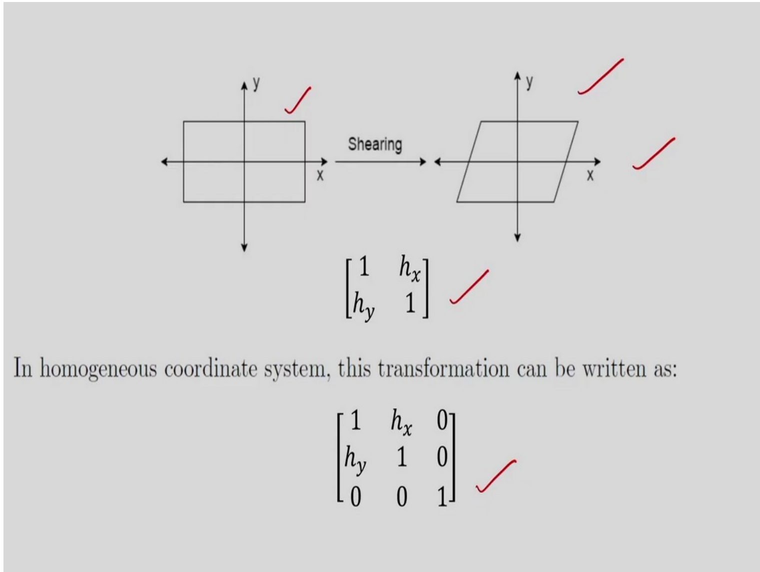
(Refer Slide Time: 07:32)



Next we consider the case of scaling of a point by a factor of Sx. Sx means scaling along the x direction and Sy along the direction of x and y respectively, Sy means scaling along the y direction. This scaling operation you can write in the matrix form like this. x' y' is equal to Sx 0 0 Sy x y, so this is the scaling operation. Now suppose if the scaling is not uniform for the whole object then it is called shearing. For example, shearing parallel to the x axis can be represented as, so x' is equal to x + ky. So, I can write like this and y'= y.

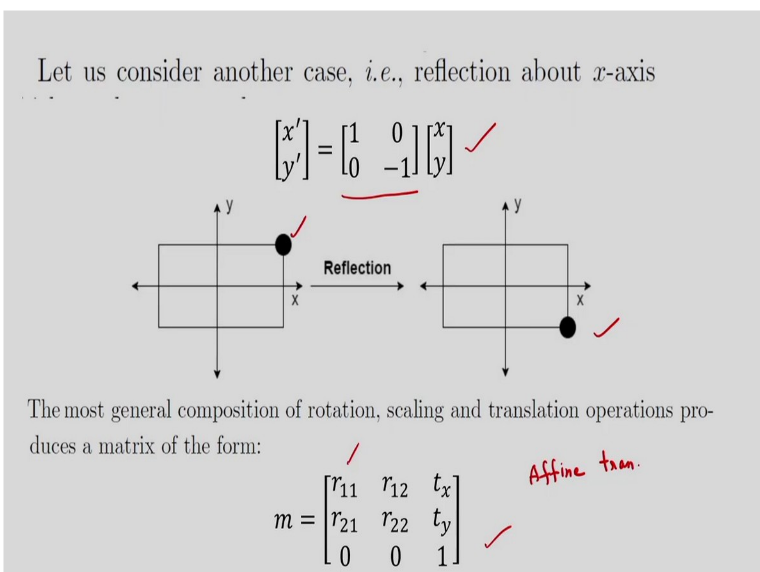
So, what is the meaning of this? The meaning is shearing parallel to the x axis can be represented as x' is equal to x + ky and y'= y. So, in the matrix form I can write this operation like this, the shearing operation I can write like this. x' y' is equal to 1 0 k 1 x y. So, this is the operation in the matrix form.

(Refer Slide Time: 8:38)



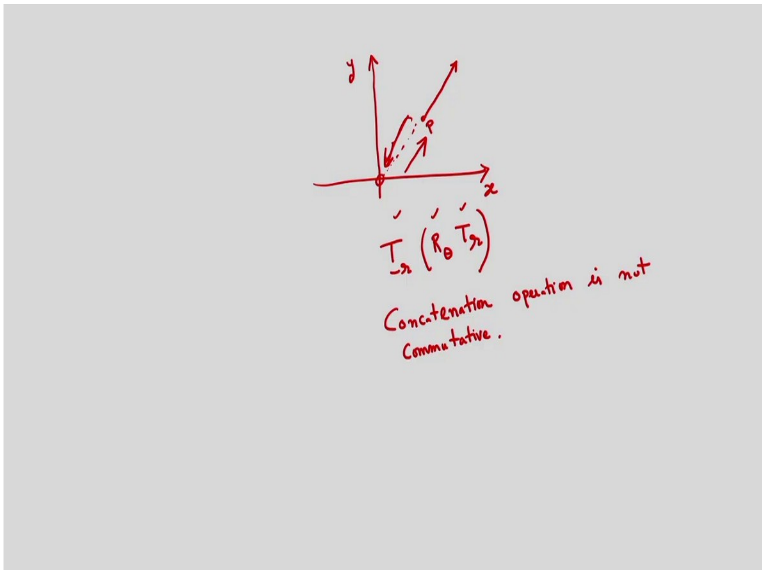
So, here I have shown the example. So, how to do the shearing? So, in this case shearing parallel to x axis I am shown here. So, this is the, suppose object and I have done the shearing. So, this is the matrix for shearing. In homogenous coordinate system this transformation I can write like this. So, the translation operation, the rotation operation, scaling operation and this operation is the shearing operations. The next I have shown one example.

(Refer Slide Time: 09:07)



This is the another case, the reflection about x axis. So, if I do this operation, x' y' is the final positions and x y is initial positions and I am considering the transformation matrix. This are the transformation matrix, 1 0 0 - 1. So, I if I do this transformation what I am getting? I am doing actually the reflection. So, if you see this one this point is reflected here. So, now the main operation is rotation, scaling and the translation and these operations I can represent in the matrix form like this. The combined operation I can show like this. The upper 2 by 2 sub-matrix is a composite rotation and scale matrix whereas, the Tx and Ty are composite translations I have considered.

(Refer Slide Time: 9:50)



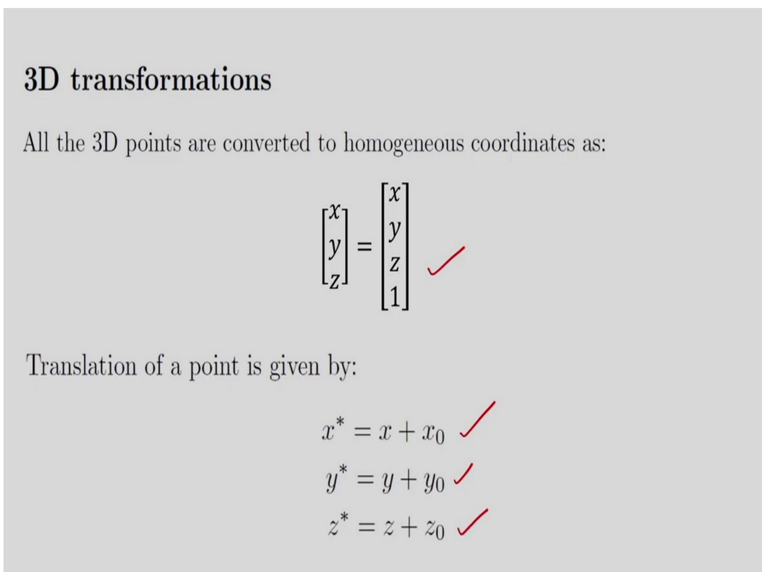
Suppose if I want to rotate a vector about an arbitrary point. Suppose arbitrary point is suppose, P. So, how to rotate this? So, for this, what I have to do? First the vector needs to be translated such that the point P is at the origin. So, first I have to do the translation operation. So, first I have to do the, this. This, this I have to move to this point, it is to be to the origin. After doing the translation operation I can do rotation operation and finally we need to translate back the point such that the point at the origin return to the point P. So, after doing the rotation again I have to do the translation back to, so that this point move to the point P. So, I am repeating this.

What is the objective? The objective is if I want to rotate a vector about a, about an arbitrary point P in 2D x y plane then in this case I have to do the following operations. First the vectors needs to be translated such that the point P is at the origin. So, I am doing the translation, so just doing the translation. So, it will be in the origin.

After this I can do rotation operation and finally we need to translate back to the point. I am doing the translation in this direction; again I am doing the translation. Translate back the point such that the point at the origin return to the point P. This complete operation I can write in this form. T - r r θ T r. Suppose I can write like this. So, what I am doing first? First I am doing the translation. After this I am doing the rotation. And after this again I am doing the translation that is in the reverse direction, - r.

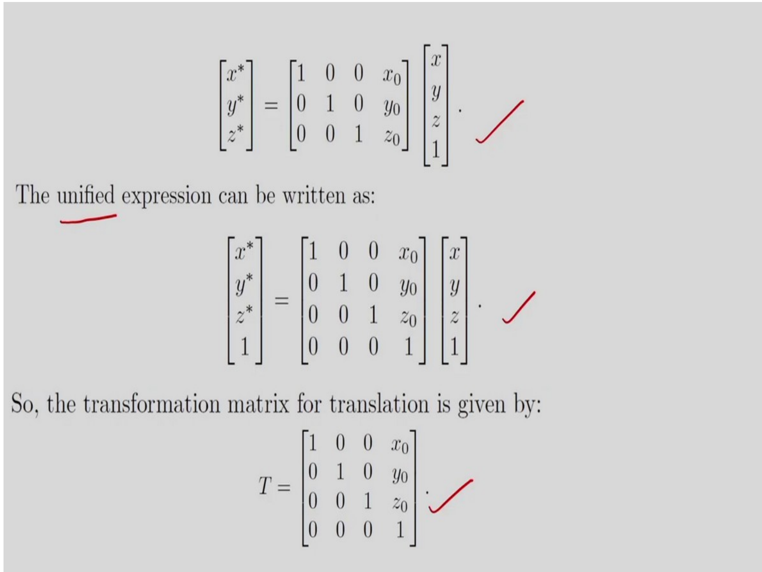
This is the combined operation I am doing. But you have to remember the concatenation operation is not commutative. That means matrix operations are not commutatives. The order of this operation is important. Suppose if I do first rotation and after this I am doing that translation again I am doing the translation, the results will be different. And one thing is important, in my previous case I have shown one, that transformation matrix. In this matrix I have considered all the operations, the rotation operation, the scaling operation and translation operation. So, this combined operation is called affine transformation.

(Refer Slide Time: 12:42)



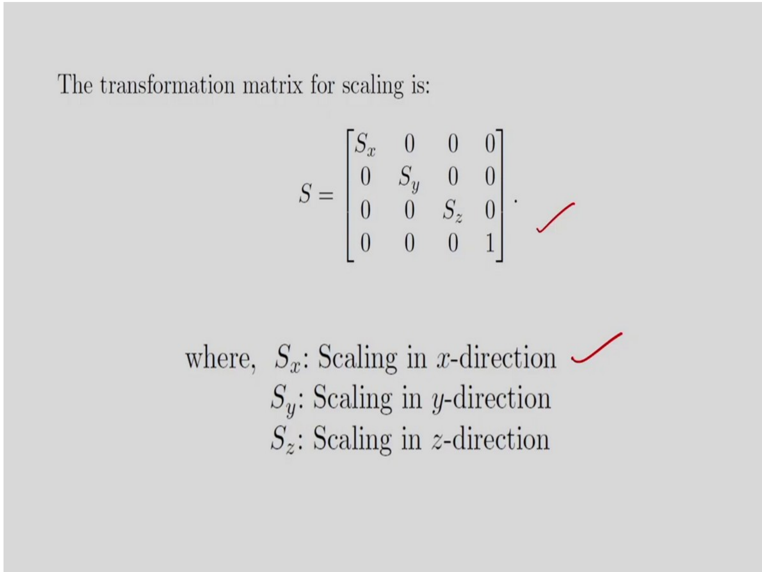
So, next I am considering 3D transformations. So, in the 3D transformation all the 3D points are converted to homogenous coordinate. It is something like this. So, x y z is converted into x y z 1. Translation of the point is given by x', $x^{¿}$= x + x_0. So, the translation along the x direction, translation along the y direction and translation along the z direction.

(Refer Slide Time: 13:11)



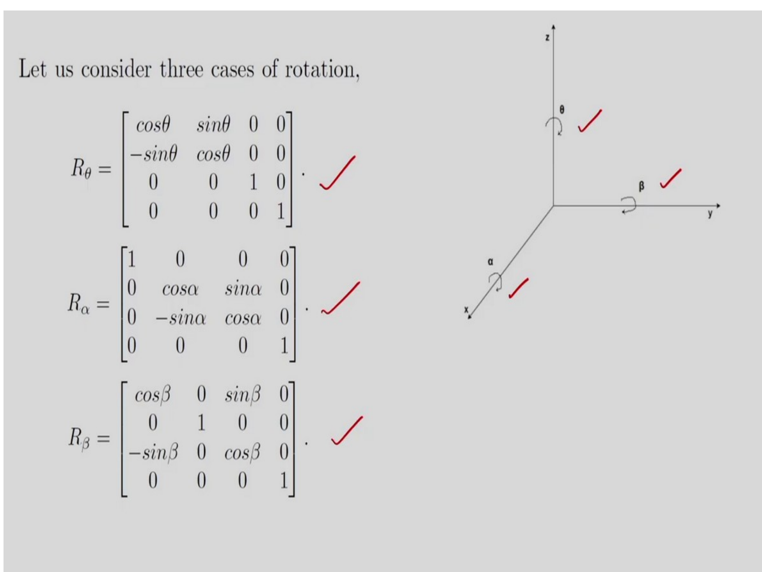
So, I can write this translation in the matrix form like this. The displacement vector is x_0 y_0 z_0. And the point x y z is translated to the point $x^{¿}$ $y^{¿}$ $z^{¿}$. And I can consider the unified expression, that is, the unified expression means the symmetric expression, so this is the symmetric expressions. So, that means the transformation matrix for translation is now converted into a square matrix. So, this is the transformation matrix for translation that is a square matrix. So, T is equal to 1 0 0 x_0, 0 1 0 y_0, 0 0 1 z_0, 0 0 0 1. So, I am considering the translation along the x direction, translation along the y direction and translation along the z direction.

(Refer Slide Time: 14:01)



Now the next I am considering the scaling operation that is the 3D scaling. The transformation for the scaling I can show you like this. So, s is equal to Sx 0 0 0, 0 Sy 0 0, 0 0 Sz 0, 0 0 01. So, this is the transformation matrix for scaling. So, in this case I am considering scaling in the x direction Sx. Sy is the scaling in y direction. And Sz is the scaling in z direction.

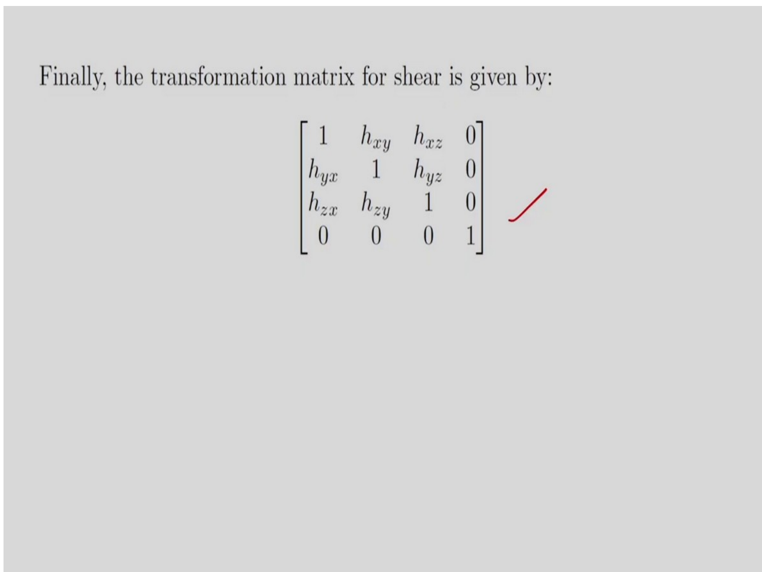
(Refer Slide Time: 14:32)



And let us consider the concept of the rotation. Now the three cases of rotations; rotation around x axis, rotation around y axis and rotation around z axis. If I consider rotation around z axis, only x and y coordinates change and z coordinates remains same. Rotation around z axis means it is a rotation in a plane which is parallel to x y plane. So, here in this figure I have shown the rotation along the x axis by an angle α, rotation along the y axis by an angle β, rotation around z axis by an angle θ.

So, corresponding to this, I have this transformation matrix. The transformation matrix like r θ that is rotation around z axis, R α is the rotation around x axis and R β is the rotation around y axis. So, I have considered these transformations.

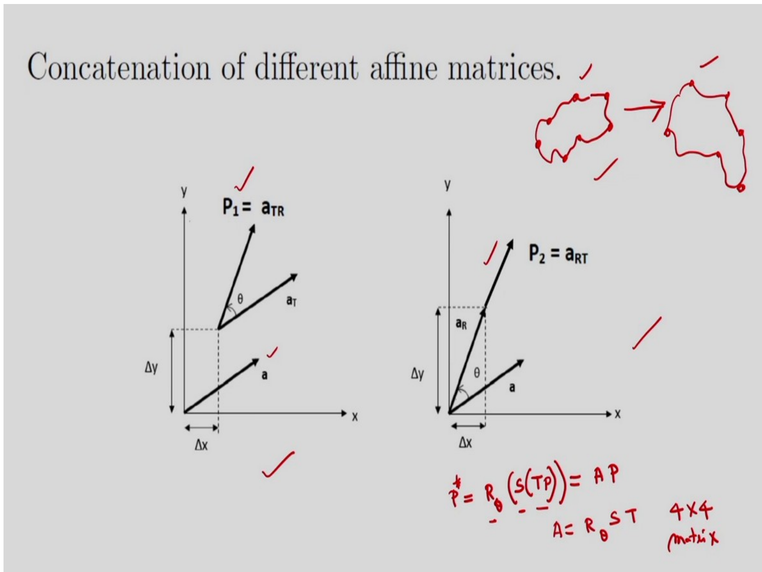
(Refer Slide Time: 15:34)



And finally, the transformation matrix for the shear is given by this matrix. So, uptil now I have discussed the concept of the transformation, I discussed the concept of 2D transformation and after this I have discussed the 3D transformation. So, in case of the 2D transformation I discussed the translation, rotation and scaling and also I discussed about shearing. In case of the 3D transformation I considered again the translation operations, rotation operations and scaling operations.

This combined operation; I can do the operations one by one so that is called the affine transformation. This is very important, affine transformation. If I understand the concept of the affine transformation then it is easy to understand the concept of image formation in the camera. So, we have number of projections like perspective projections, orthographic projections. So, for these projections I have to understand this concept, the concept of the homogenous coordinate systems. Now already I had discussed that the matrix operation is not commutative, the order of this operation is important.

(Refer Slide Time: 16:47)



So, I can give one example in this case, so if you see the next slide, so in this case what I am doing? In the first figure, in this figure I am considering one vector. The vector is a. First I am doing the translation of this vector and after this I am doing the rotation of this vector by an angle θ. So, corresponding to this I am getting the point is P1. So, the meaning is, first I am doing the translation of the vector by an amount to Δ x in the x direction, Δ y in the y direction and after this I am doing the rotation of this vector.

In the second case what I am doing? First I am doing the rotation of the vector, the same vector. And after this I am doing the translation. So, I am getting P2. So, it can be observed that the P1 is not equal to P2. Because the matrix operation is not commutative the order of operation is quite important. And in case of the affine transformation suppose, suppose if I do something like the translation I am doing and after this I am doing the scaling and after this if I do the rotations, so how can you write this one?

So, suppose I am getting one point, new point, P, this one and R θ, S T P is equal to A P. So, in this case what I am considering, this will be T P, so in this case what I am doing? In this case first I am doing the translation of this point T P. After this I am doing the scaling, S is the spelling and after this I am doing the rotation. So, this combined operation is called the affine transformation. So, the matrix, what is the matrix, A, the matrix A is equal to R θ S and T. So, this matrix will be 4 by 4 matrix.

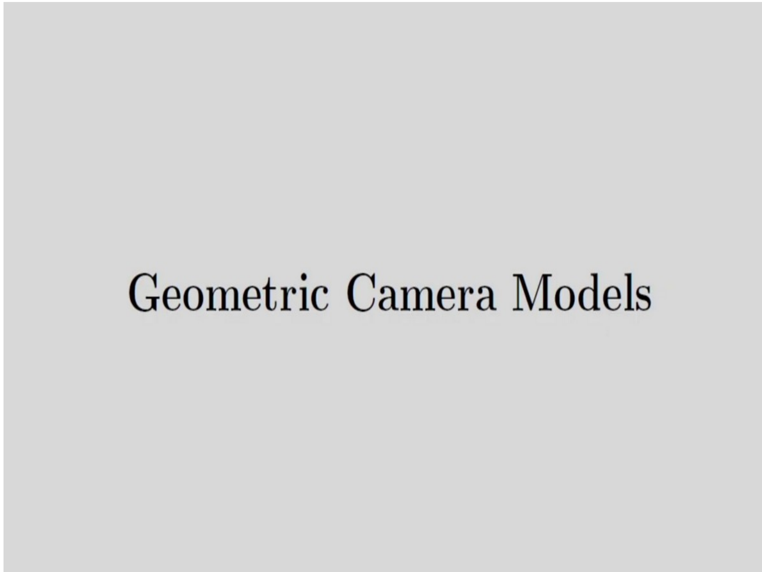
So, this matrix I can consider as affine matrix. So, first I am doing the translation. After this I am doing the scaling. And after this I am doing the rotations. So, a is equal to R theta S T. In the first case I have shown I have done the translation and after this I am doing the rotation. In the second case what I have done? First I have done the rotations and after this I am doing the translation. In this case the P1 is not equal to P2.

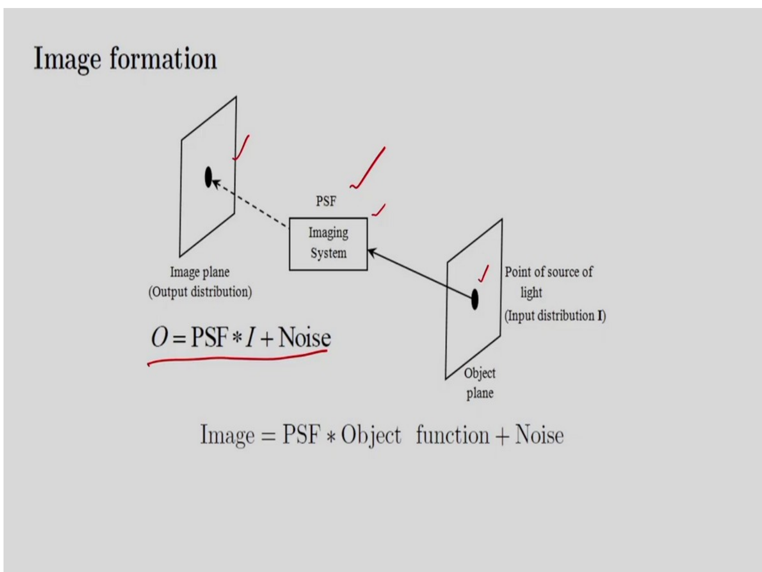
So, these matrix operations are quite important. So, I can give one example, suppose. I have some objects. This object is like this in one image. In another image suppose object will be like this. So, now I have to determine what type of transformation it is going on. It may be rotated. The object may be rotated. It may be translated and some scaling operations may be done.

So, corresponding to this you have to determine the transformation matrix. So, for this, suppose some of the keypoints you can consider, and based on these keypoints you can understand what type of transformation it is going on that is, you can find the transformation matrix, whether it is rotated, whether it is scaled or the translation operation is going on. So, I have given this example.

So, in an image, suppose this is the first image, in the second image I have this one. So, what type of transformation it is going on? So, that you can determine. So, this is about the geometric transformation. So, these transformations are important for understanding the concept of image formation in the camera.

(Refer Slide Time: 20:33)



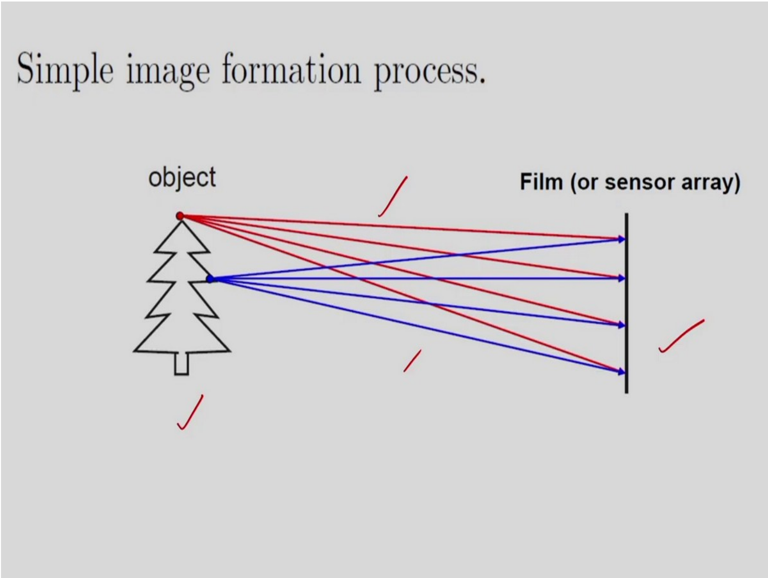


Now after this, I will discuss the concept of the image formation in the camera. So, if you see in my next slide I am going to discuss about the geometric camera models. So, first one is image formation principle. And in my last class I explained about this, what is the output image? The output image is the PSF. And it is convolved with the object function + noise, noise of the imaging system.

The object function is a object or a scene that is being image. And what is the PSF? The PSF, the Point Spread Function is the impulse response when the inputs and the output are the intensity of light in an imaging system. That means, it represents the response of the system to a point source. So, PSF indicates the spreading of the object function. And it is the characteristics of the imaging system. So, PSF is quite important.

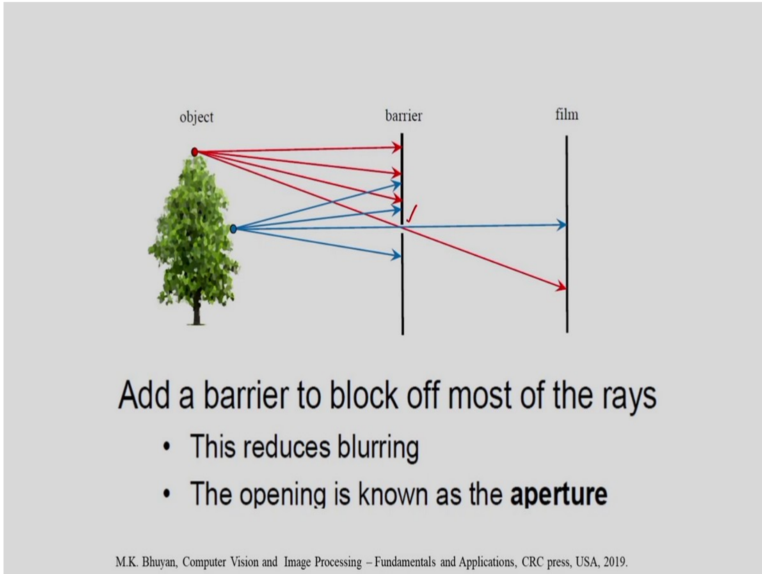
So, this is my object and I am getting the image here. So, corresponding to this here, you can see the PSF, it is convolved with the image and I am considering the noise of the imaging system. And one important point in the last class I discussed, a good or the sharp imaging system generally has a narrow PSF, whereas a poor imaging system has a broad PSF.

(Refer Slide Time: 21:46)



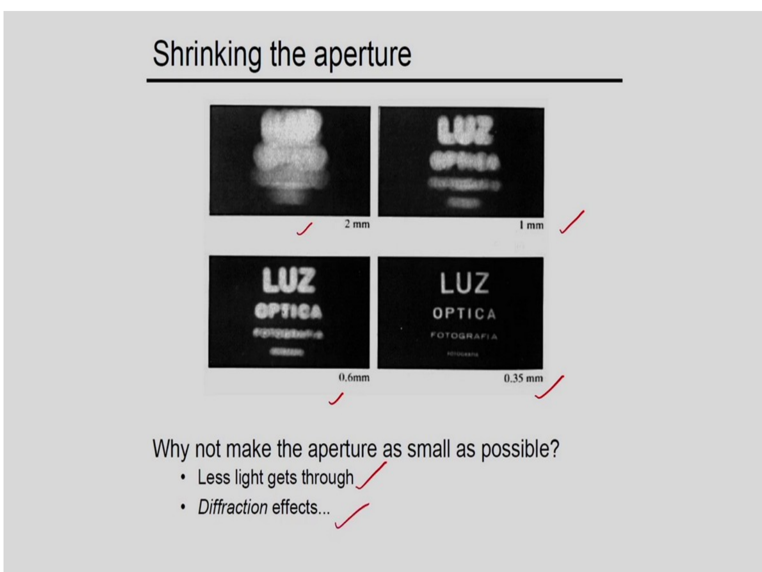
And here in this case I have shown simple image formation process. So, I have shown one object and I have shown the film that is the sensor I am considering. The sensor converts the light photon into electrical signal. But is this case if I consider, I will be getting the blurred image. Because it is the (inter) intersections of all the light rays if I consider. So, because of this I am not getting the fine quality image. I am getting the blurred type of image. So, that is why I can consider this setup.

(Refer Slide Time: 22:21)



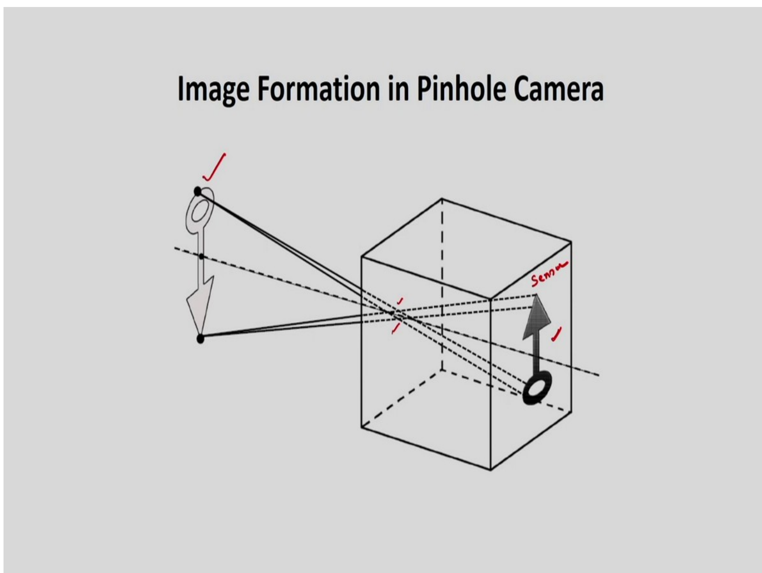
That means I am considering one barrier. And in the barrier, one aperture is there, small opening is there, this opening is there. If I consider this opening, that means a barrier is considered to block most of the rays. So, because of this structure it reduces blurring. In my earlier slide I have shown I am not getting the image, the good quality image because of the number of rays. And in this case it is incident on the sensor. So, in the second case I am getting the good quality image because it reduces blurring. And this opening is called the aperture, the aperture of the imaging system.

(Refer Slide Time: 22:58)



So, suppose if I consider the shrinking the aperture. Suppose if I want to reduce the size of the aperture what will happen? So, in the first case I am considering the aperture size is 2 mm, next is 1 mm, after this 0.6 mm, next is 0.35 mm. That means gradually and gradually I am I am reducing the size of the apertures, shrinking the apertures. So, in this case I am getting sharp images, but we cannot make the aperture as small as possible. Because if I make the aperture very small less light gets through. And also I have to consider the effect of diffraction. That is why I cannot make the aperture as small as possible.

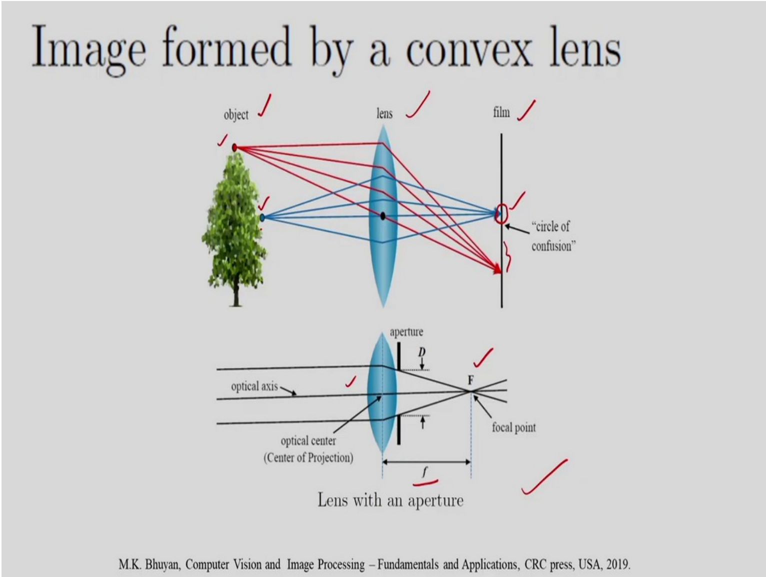
(Refer Slide Time: 23:38)



And in this case I have shown the image formation in a pinhole camera. So, this is a very simple camera. So, I have one opening here, small opening, hole is there. And this is my object, and this is the image.

So, I am getting the inverted image here. So, this structure is very much similar to human eye. In case of the human eye, this, instead of this opening, I have pupils and I have the retina. In case of the camera I have the sensor here. In case of the human eye we have retina. In retina also we have two types of photoreceptors, the rods and the cons. Rods are responsible for monochromatic visions and cons are responsible for color visions. So, this structure is very similar to human eye and I am getting inverted image here.

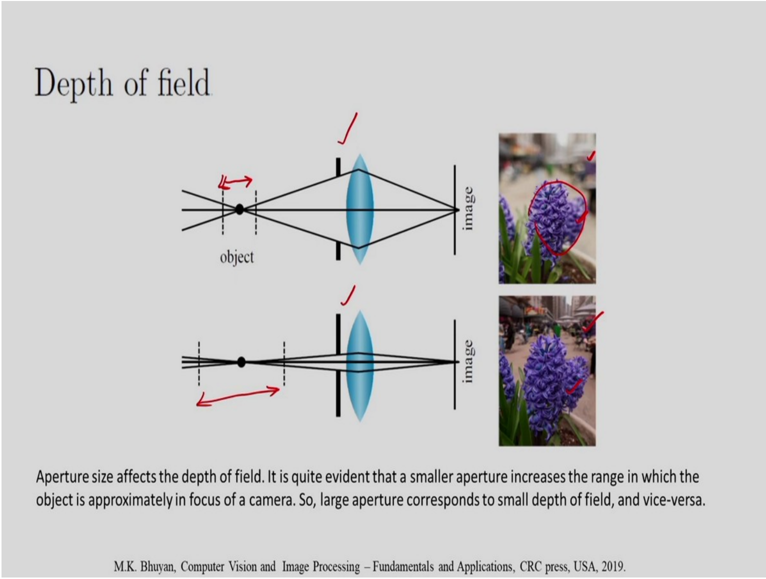
(Refer Slide Time: 24:25)



And in this case I have shown the image formed by a convex lens. So, I have the objects and I have lens. This is the convex lens. I have the sensors film. So, if I consider this portion here, this portion of the object, this is properly focused. So, that means, in this case I am getting the image here. And in this case, this portion if I consider, this portion of the object; that is not properly focused. Then in this case it is nothing but the circle of confusion. Corresponding to this point, this point is properly focused by the lens, the convex lens. So, I am getting the image here. This portion I am getting the image. But corresponding to this portion, this is not properly focused. So, I am getting the blurred portion here. And this is called the circle of, circle of confusion.

And in the second diagram I have considered lens with an apertures. So, small aperture I have considered. Because why, what is the need of the aperture, I have already discussed. So, aperture is important. this is optical centre and the centre of projection. This is the optical axis. And after this, this is the focal point in the convex lens. This is the focal length and I have shown the image formation process in the convex lens.

(Refer Slide Time: 25:40)

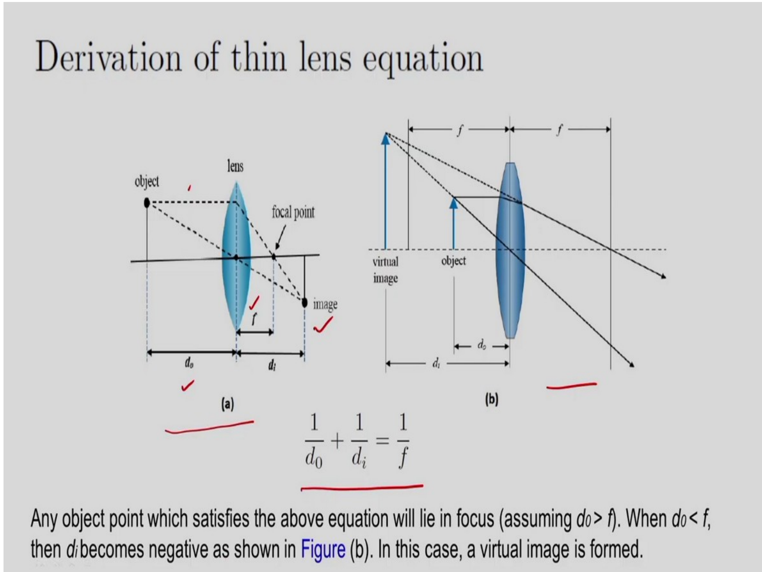


And there is a concept of the depth of field. So, here you see I have shown two cases. In this, in the first case I have shown this portion, suppose if I consider the aperture is wide. And in this case the aperture is narrow. In the first case the aperture is wide, in the second case aperture is narrow. In the first case, only this portion of the object, from this length to this length it is properly focused, the rest of the portion, it is not properly focused.

So, corresponding to this portion I am getting the sharp image. And if I consider the background I am not getting the sharp image, I am getting the blurred background. That means this portion is properly focused. That means corresponding to this depth this is properly focused. In the second case, the aperture is narrow aperture so that means this range I can expect, that this range it is properly focused. So, that is why this is coming, this is coming properly. And the background is also coming properly.

That means the aperture size affects the depth of field. And it is quite evident that a smaller aperture increases the range in which the object is approximately in focus of a camera. So, large aperture corresponds to a small depth of field and vice versa.

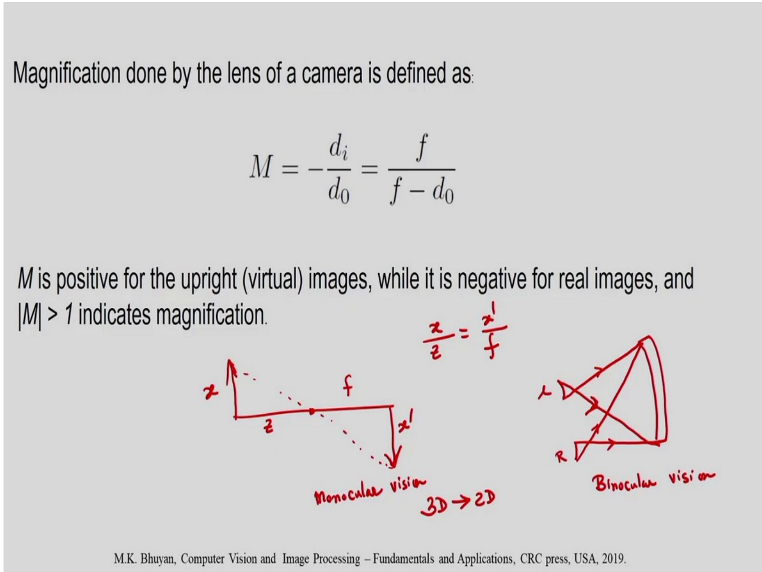
(Refer Slide Time: 26:55)



And in this case I have shown the derivation of the thin lens equation. I am considering the convex lens and from the similar triangles, this equation I am getting. What is d 0? d 0 is the distance between the objects and the lens. So, this is d 0 What is di? The distance between the lens and the image. So, I am having the image here. And here I have shown the focal point and also the focal length I have shown. f is the focal length.

So, in the first example I am getting the real image. In the second case I am getting the virtual image. So, if I assume d 0 is greater than f, then in this case then object point which satisfies the above equation will lie in focus if I consider d 0 is greater than f. And when d 0 is less than f, then di becomes negative. Then in this case I will be getting the virtual image. So, I have shown this derivation of the thin lens equation. Already you know these equations.

(Refer Slide Time: 27:54)



And what is the magnification done by the lens of a camera? The magnification is defined by di divided by d 0. That is the definition of the magnification. M is positive for the virtual image while it is negative for the real image. And if I consider M is greater than 1 that indicates the magnification. So, this is the definition of the magnification done by lens of a camera. So, in case of the monocular setup, monocular vision, suppose this is the object and this is the image.

So, this is my, suppose, x and this is x', suppose and this is z and suppose this is f. So, this is the monocular vision setup. So, corresponding to this, you know this equation from the similar triangle x / z $= x'$ / f, monocular vision. And if I considered suppose two cameras. I am considering one surface something like this. I am considering one camera another camera is this.

So this camera, suppose consider this one there and this I am considering. So, in this case I am considering two cameras. So, I can consider this as left camera, suppose and this is right camera, suppose. So, this setup is called binocular vision. Binocular vision. So, in the monocular vision because image acquisition system is nothing but the 3D to 2D projections so I do not have the depth information in this case. But with the binocular vision I can get that depth information.

So, in my next class I will discuss the geometrical camera models like the some projection models I am going to discuss, like the perspective projections, orthographic projections. And after this I will discuss about the stereo visions. So, by using the stereo vision you can get the depth information from disparity. So, that concept I am going to discuss in my next class. So, today I have discussed about the geometry transformation which is very important like the affine transformation. And based on this affine transformation, you can understand the concept of the perspective projection and the orthographic projection. And you remember the concept of the homogenous co-ordinate systems. So, that is all for today. Let me stop here today. Thank you.