Computer Vision and Image Processing – Fundamentals and Applications

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

Lecture – 08

Image Formation in a Stereo Vision Setup

Welcome to NPTEL MOOCS course on Computer Vision and Image Processing Fundamentals and Application. Up till now, I discussed about the image formation concept in a single camera. So, in a single camera setup there are many disadvantages. So, one major disadvantage is that in single camera based setup it has limited field of view. For some applications like object detection or maybe object tracking we need large field of view.

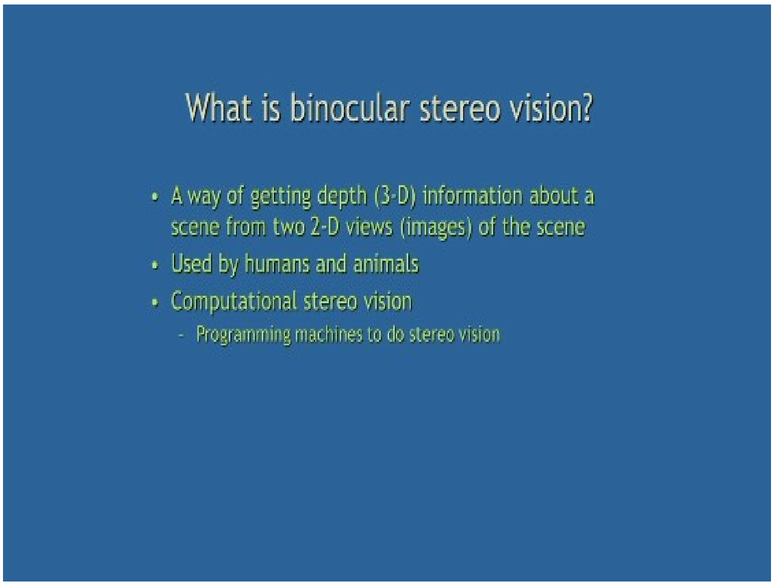
So, in a single camera based setup it is not possible to get the wide field of view the large field of view. So, that is why we have to use multiple cameras instead of single camera. And another problem is you know that in image formation system it is nothing, but the 3D to 2D projection x, y, z coordinate is projected into x, y coordinate and because of this the depth information is lost.

So this is the concept of the image formation. Now because of this if I consider only single camera it is very difficult to get the depth information. I have discussed some techniques to get the depth information, the shape information. Some cubes like shape from shading, I discussed about this shape from shading. Shading means the variable levels of darkness and after this the shape from textures I have highlighted this concept.

Shape from motion or maybe the shape from focus and defocus images and shape from contours. Depth estimation from a single camera is an ill-posed problem. So, for depth estimation we need minimum two cameras. In a stereo vision setup that is the binocular setup we have two cameras. In two cameras we have two images. One is the left image another one is the right image.

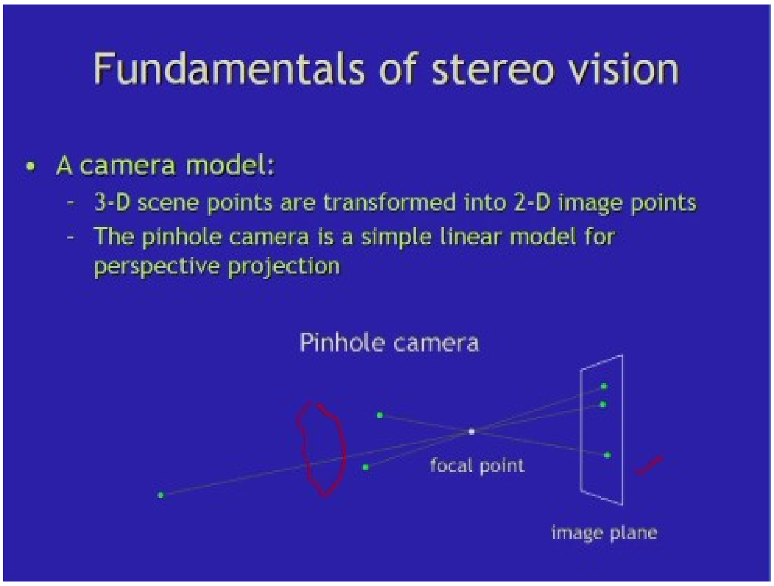
From this two images I can get the depth information. So, today I am going to discuss about the image formation principle in a stereo vision setup.

(Refer Slide Time: 02:25)



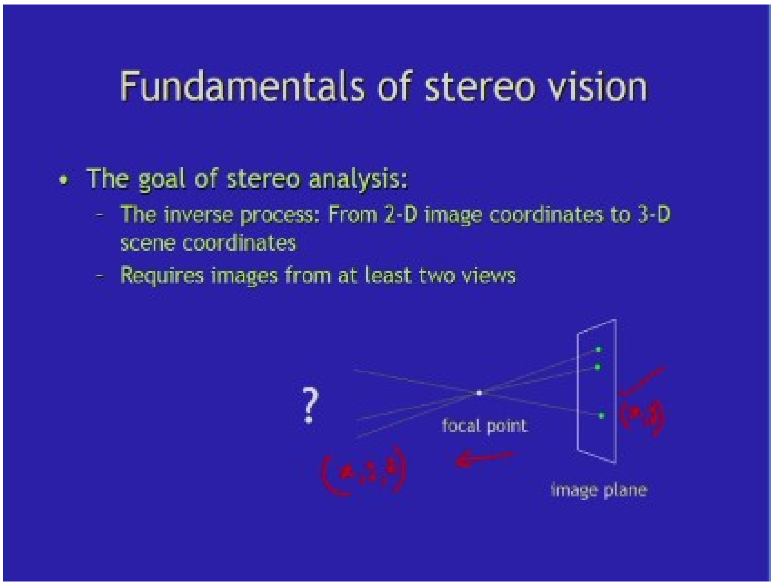
So, let us see what is the stereo vision setup? Now already I told you that image formation process is nothing, but the 3D to 2D transformation. The 3D scene coordinate is transformed into a 2D image coordinates. In case of the human visual system, human brain reconstruct the 3D information. In case of the computer vision we have to develop some algorithm so that computer can determine the depth information. So, for this we can consider a stereo vision setup.

(Refer Slide Time: 02:58)



So, here I have shown the image formation concept that I have already explained that is the pinhole camera model that is the very simple model and simple linear model and it corresponds to the perspective projection. So, I have shown the pinhole camera and I have shown the image plane, the image plane is shown and also the pinhole camera and suppose if I consider this is the object. So, object maybe something like this, so this is the object. I am getting the 2D image in the image plane. So, this is nothing, but the perspective projection.

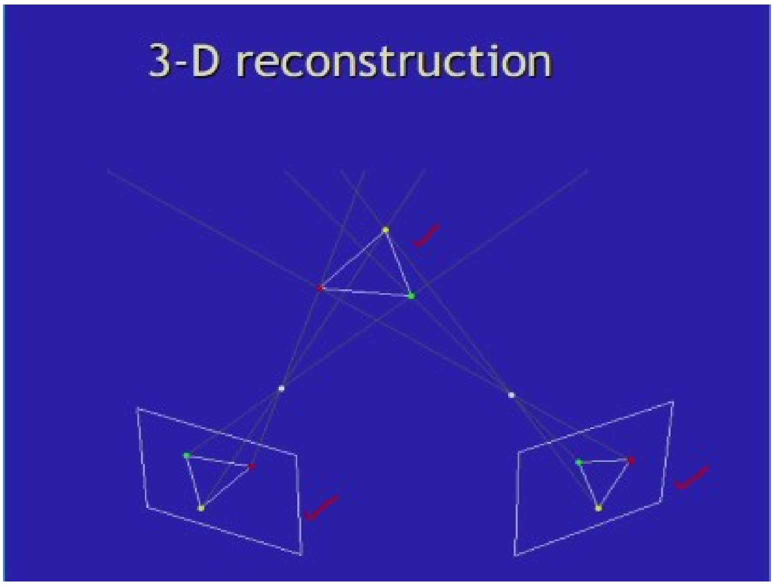
(Refer Slide Time: 03:31)



Now in case of the stereo vision, the goal of the stereo analysis, we have the 2D image points here. We have the 2D coordinates, the 2D coordinate I can write like this x, y so this is the 2D coordinate x, y and from the 2D coordinate I have to get the 3D coordinate of the scene the 3D coordinate is x, y, z. So, I have to determine this.

This is the objective of the stereo analysis the goal of stereo analysis. From the 2D image coordinate, I have to get the 3D scene coordinates. And in this case I need minimum two images for this estimation the depth estimation the 3D information.

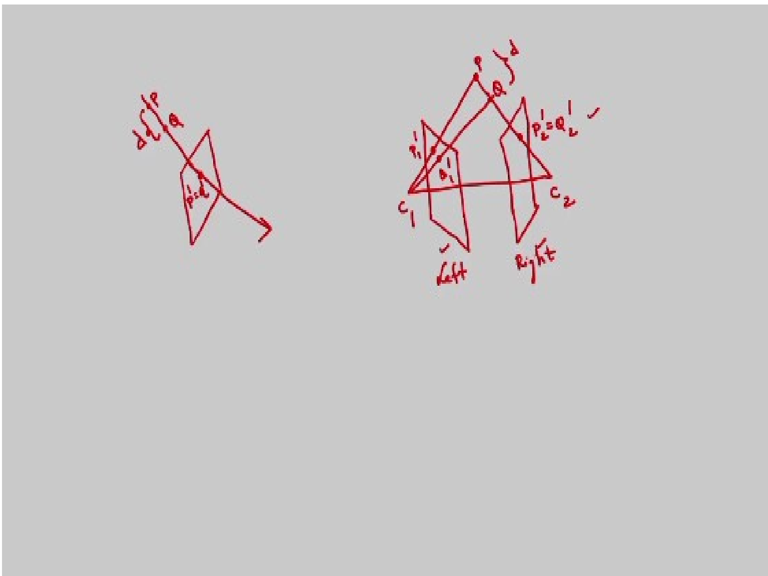
(Refer Slide Time: 04:21)



Here I have shown the 3D reconstruction so I have shown the left image and this is the right image. From these two images, I am determining the 3D information this is the 3D information. So, already I told you that we need minimum two images. So, in this case I have consider binocular setup.

So, in the binocular image setup I have two images one is the left image another one is the right image. From these two images I am getting the 3D information the depth information I can determine.

(Refer Slide Time: 04:53)



So I can show suppose I consider one diagram here so suppose this is the image plane and I have the points so suppose this point is P and this point is Q. So, I am just doing the projection corresponding to this projection in the image plane I am getting the projection P dash is equal to Q dash same point I am getting so I am doing the projection here. Same point I am getting in the image plane corresponding to this P and Q. So, this depth information is lost here.

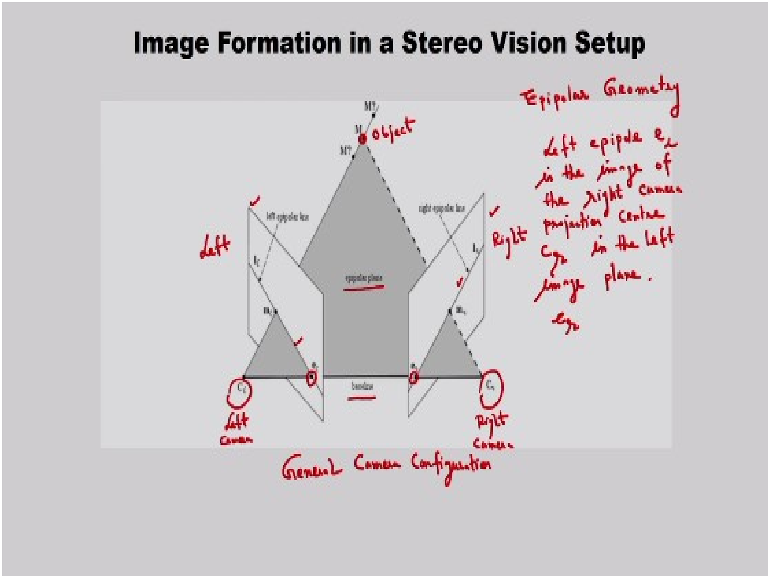
The depth information is d depth information is lost in this case, but if I consider one stereo vision setup suppose if I consider one stereo vision setup something like this.. So, I am considering two points P and Q corresponding to this P and Q I have one projection that is the projection is P1 dash or maybe P1 dash and corresponding to the point Q this point is Q so I have the another projection this is Q1 dash.

I am considering one camera is C1 another camera is C2 that is the center of projection and corresponding to the second camera, second camera is C2 so I have the projection here. This projection I am getting the same point so it is P2 dash is equal to Q2 dash same point I am getting for about the points. So, this is the depth information. So, in the second setup that is the stereo vision setup, there is a binocular setup.

You can see you can see in one image plane I am getting this projection in second image plane I am getting this projection. So, from this information I want to get the depth information. So, I can show how to get the depth information from these two images one is the left image I can consider is a left image and this is the right image. So, C1 and C2 are mainly the camera projection center C1 is the left camera and C2 is the right camera.

So, in this setup you can see in the binocular setup in the stereo vision setup I have two images and you can see the projections of the point P and Q corresponding to the left image and corresponding to the right image. So, corresponding to the right image I have the same point the P2 dash is equal to Q2 dash, but considering to the left image the projection points are different. So from this information I want to get the depth information. So that is the fundamental concept of the stereo vision.

(Refer Slide Time: 08:14)



So in this setup I have shown the image formation in a stereo vision setup. So this configuration is called the epipolar geometry this is called the epipolar geometry. So, what is the definition of the epipolar planes? So, here you see I have the left camera projection center, this is the left camera projection center, this is the right camera projection center and corresponding to this two cameras I have the image plane.

One is the left image plane this is the left image plane and this is the right image plane. So, I have two image planes and I am considering the point the point is suppose M that is the scene coordinate that is the 3D point I am considering the point is M. So, first I have to define the epipolar plane. The epipolar plane is nothing, but the plane joining the points the camera projection centers and the 3D point the 3D point is M that plane is called the epipolar plane.

This plane is called the epipolar plane. The plane corresponding to the camera projection centers the left camera and the right camera and also the 3D point the 3D point is M and also I have the image planes. So, I have two image plane one is the left image plane I have shown already this is the left image plane and this is the right image plane. So now I want to define the epipolar lines. So, I have two epipolar lines.

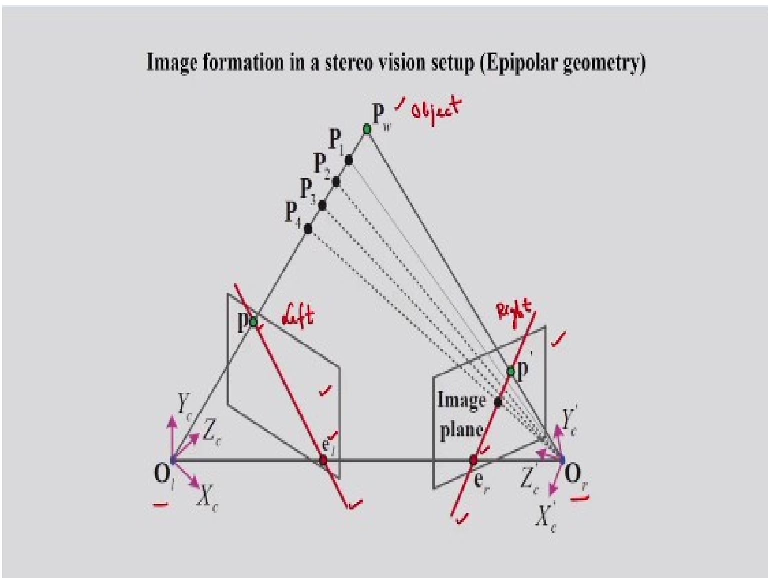
One is the left epipolar lines this is the left epipolar line you can see and one is the right epipolar line. So, epipolar lines means the intersection of image plane with epipolar plane you can see the epipolar lines means this is the intersection of image plane with epipolar planes that is the definition of the epipolar lines and corresponding to this I have the baseline if you see the baseline between the camera.

So, I have the left camera this is my left camera and this is my right camera, this is the left camera projection center. So the line between the left camera projection center and the right camera projection center that is called the baseline. So, this line is called the baseline. So, I have two epipoles one is the right epipole this is the right epipole another one is the left epipole this is the left epipole.

Epipole means the intersection of image plane with the baseline. So, intersection of the image plane with the baseline that is called the epipoles. So, I have two epipoles one is the left epipole another one is the right epipole. So, I can write the left epipole so left epipole is el in this diagram left epipole el is the image of the image of the right camera projection center, right camera projection center is C2.

Here, I am writing cr right camera projection is cr in the left image plane. So, left epipole el is the image of the right camera projection center cr in the left image plane also I can define the right epipole, right epipole is er. So, I can define the right epipole. So, this camera configuration is called general camera configuration. This is the stereo vision setup and this configuration is called general camera configuration. In this case you can see the image planes are not parallel.

(Refer Slide Time: 12:35)



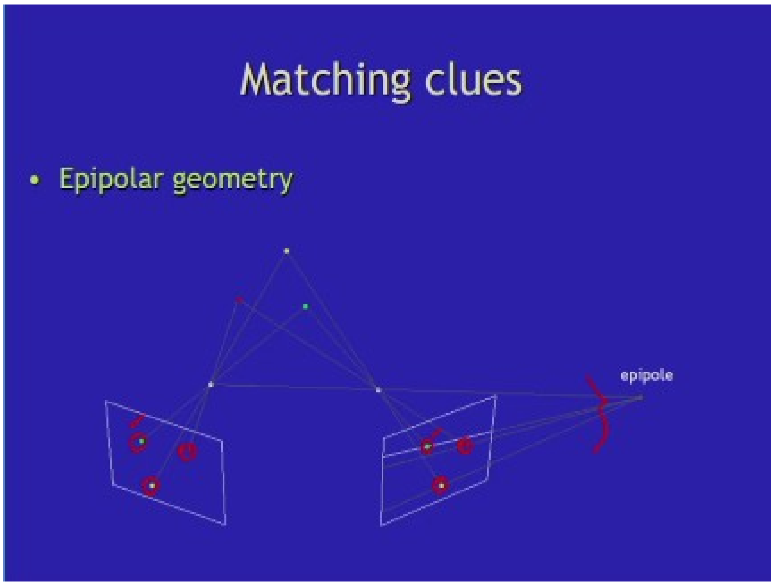
Next, we can see the same configuration I have shown here the epipolar geometry I have shown. So this points P1, P2, P3, P4, Pw these are the points that is the object points in the world these are the objects basically object points in the world and I have the left camera projection center Ol and the right camera projection center Or and already I have defined the epipolar plane.

The plane corresponding to the points Ol, Or and Pw that is the epipolar planes. I have two image planes, one is the left image plane and on the right image plane and I have shown the projection of the points. So, corresponding to the point this point P1, P2, P3, P4, Pw I have only one projection in the left image plane so this is my left image plane. So, I have only one point here.

And corresponding to the right image plane so in the right image plane you can see the projection of the points the points are P1, P2, P3, P4. So I have change the projection of the points in the image planes right image plane and the left image plane and in this case I have shown the epipolar lines. One is the left epipolar lines, this is the left epipolar lines and this is the right epipolar lines and also I have shown the epipoles.

One is the left epipole el another one is the right epipole that is er is the right epipoles. So, you can see that the image formation in a stereo vision setup and I have shown the two images one is the left image another one is the right image. Corresponding to the left image, I have shown the projection of the points P1, P2, P3, P4 and Pw and corresponding to the right image plane I have shown the projection of this points P1, P2, P3, P4. So from these two images I want to get the depth information.

(Refer Slide Time: 14:31)

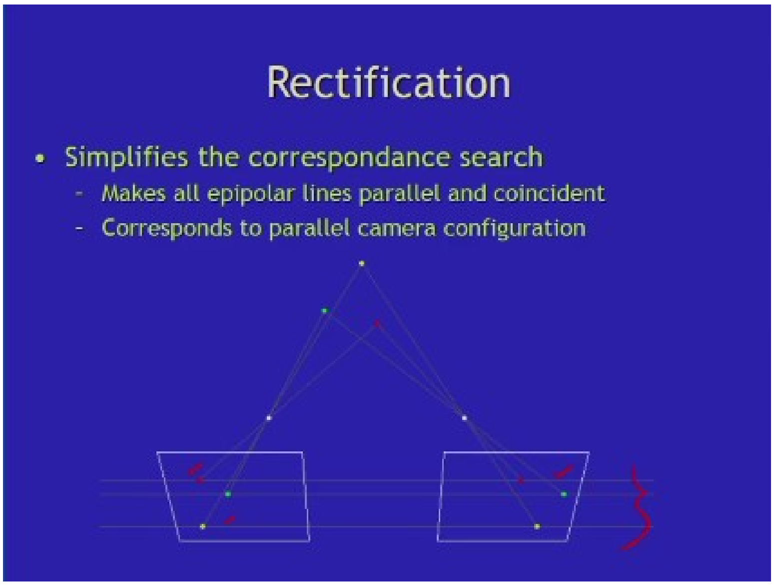


So, in this case for getting the depth information I have to match the points one of the points in the left image another one is the points in the right image. So, here I have shown in this diagram the matching of the points. So, here the green point is matched with the green point here one is the left image another one is the right image. Similarly, that yellow point is also matched the yellow point is also matched and the red point is also matched.

So, I have to do the matching so after the matching I can find the correspondence between these two images. One is the left image another one is the right image and in this case you can see that epipolar lines are not parallel in this case you can see this epipolar lines these are the epipolar lines the epipolar lines are not parallel and in this case I can show you that if I want to find the correspondence between the points I have to search along the horizontal direction along the vertical directions.

Suppose, if I want to find the correspondence between this point and the green point so that means I have to do the searching. So, searching along the horizontal direction and searching along the vertical direction in two directions I have to search the points to find the correspondence between the images the two images. One is the left image another one is the right image.

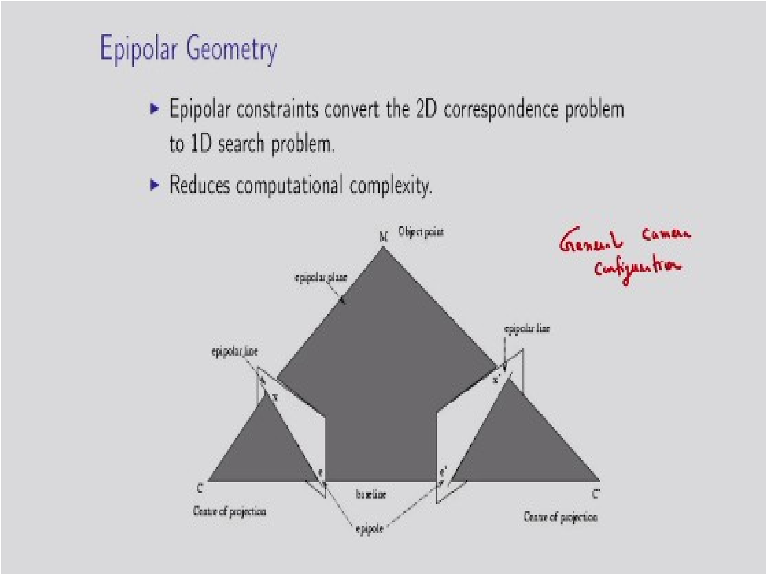
(Refer Slide Time: 15:51)



But in this configuration I have shown the epipolar lines are parallel now. So, I am making the epipolar lines parallel and actually this configuration is called the canonical stereo configuration. So, in this case you can see the 2D searching problem is converted into 1D searching problem.

So, along the epipolar lines you have to search. The red point is searched along this line, the green point is searched along this epipolar line. The yellow point is searched along the epipolar lines that means the 2D searching is converted into 1D searching. This is called rectification the image rectifications.

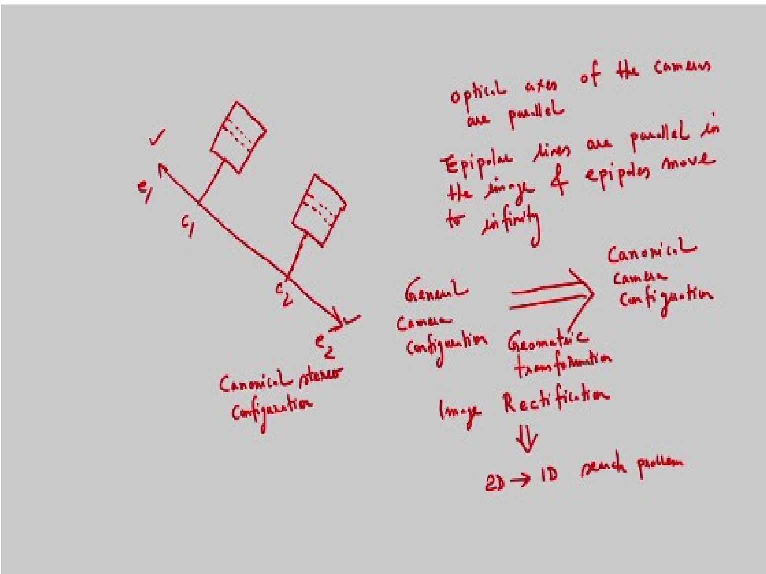
(Refer Slide Time: 16:31)



So, here I have shown the same concept. The 2D correspondence problem is converted into 1D search problem that is it reduces computational complexity. In this configuration, this is the general camera configuration already I told you this is the general camera configuration. So, in this case the epipolar lines are not parallel.

So, in this case if I want to find the correspondence between the left image and the right image. I have the search along the horizontal direction and along the vertical directions that is 2D searching I have to do.

(Refer Slide Time: 17:09)



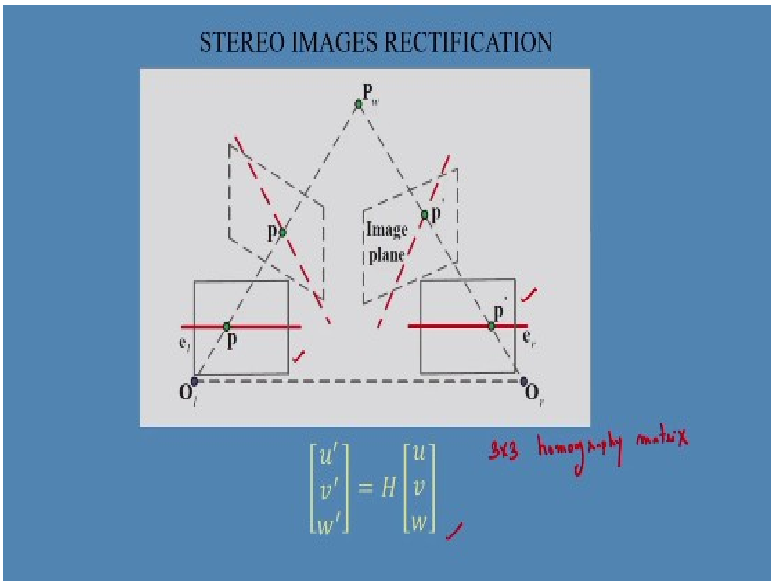
Suppose, I am considering this configuration suppose so this is the camera projection center C1. Two image planes I have shown here, so this is also C2 and epipoles I am showing e1 and e2. So, this configuration is called canonical stereo configuration. So, in this configuration optical axes of the cameras are parallel and epipolar lines are parallel in the image and epipoles moves to infinity.

You can see the epipole moves to infinity I have two epipoles e1 and e2. So epipole moves to infinity and in this case the baseline is aligned to the horizontal coordinate axes. So, this configuration is called the canonical stereo configuration. In this case, the computation is simpler I can convert the general camera configuration that already I have defined the general camera configuration.

I can convert the general camera configuration into the canonical camera configuration. By using some transformation this geometric transformation I have to do, so I have to do some geometric transformation I can convert the general camera configuration into canonical camera configuration this is called the image rectification. So thus image rectification actually it converts the 2D the search problem from 2D to 1D search problem.

So, I have shown that how to convert the general camera configuration to canonical camera configuration by using some geometric transformation and that concept is called the image rectification. So after doing the image rectification, the epipolar lines will be parallel and the epipoles move to infinity then in case the 2D search problem will be converted into 1D search problem. So, search is important to find the correspondence between the images the left image and the right image.

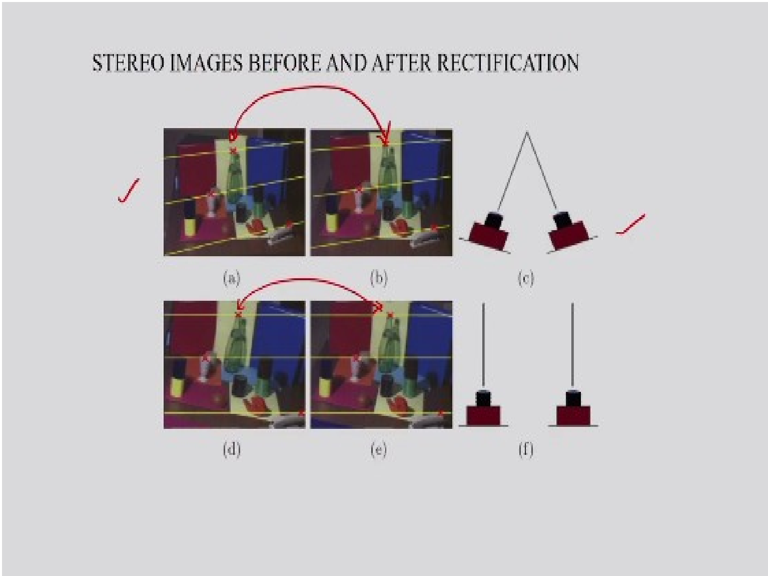
(Refer Slide Time: 21:33)



So, in this case I have shown that one is the general camera configuration. In the general camera configurations, epipolar lines are not parallel, but in the second case if you can see here the second case I am considering the canonical configuration. In the canonical configurations, the epipolar lines are parallel. Now, in this case I have to do some transformation to convert the general camera configuration into canonical configuration.

Here I have to show what, I have to do, so I have considered the 3 by 3 the homography matrix, so I have considered the 3 by 3 the homography matrix, I have considered. So, I am doing I am doing the transformation the transforming the coordinate of the original image plane I can do by using this transformation. So for this I am considering the 3 by 3 homography matrix.

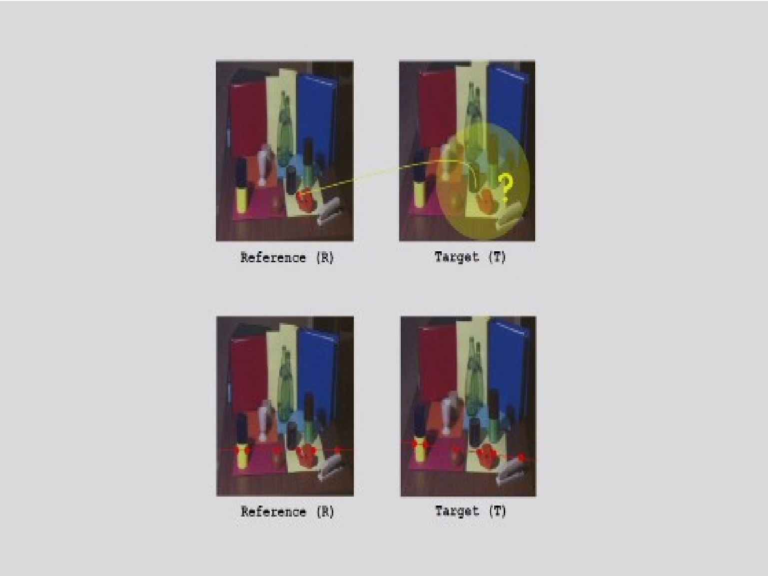
(Refer Slide Time: 22:46)



And in this diagram I have shown that the stereo images before and after rectifications. So, in the first case if I see this first case the epipolar lines are not parallel and corresponding to this you can see the image planes these are also not parallel, but in case of the second case after the rectification epipolar lines will be parallel. So, in the first case if I want to find the correspondence between the images I have to search along the horizontal lines and the vertical lines in two directions I have to search.

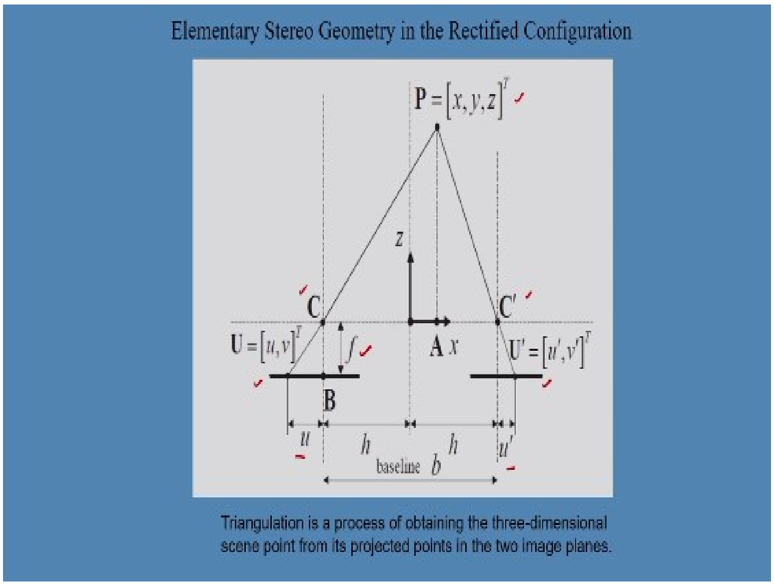
So, suppose if I want to find the correspondence between these two points I have to search the image, but in the second case if I want to find the correspondence between the points I have to search along the line particular line so that line is the epipolar lines. So, that means the 2D searching is converting into 1D searching because of the image rectification.

(Refer Slide Time: 23:39)



Here again I have shown another example that is before rectification and after rectifications. So, how to find the correspondence between the two images, one is the reference image another one is the target image.

(Refer Slide Time: 23:51)

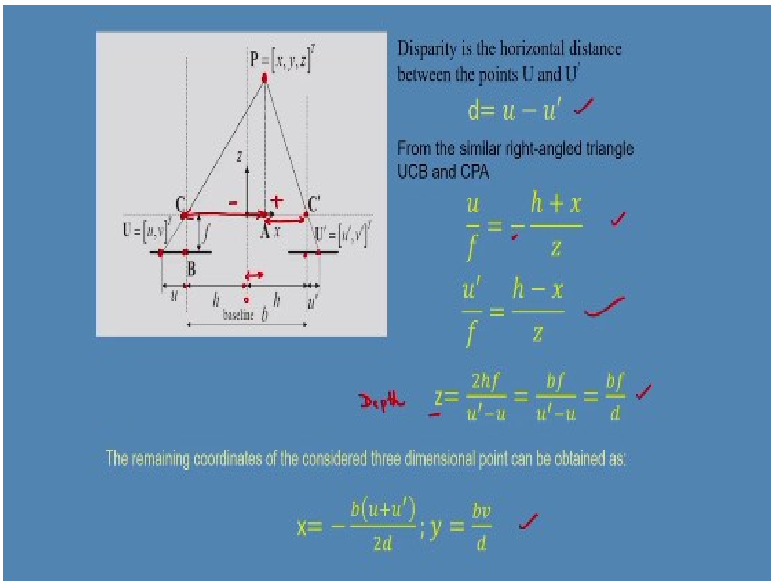


Now, here I have shown the elementary stereo geometry in the rectified configuration. So that means from this configuration I want to determine the depth information. So, let us see this configuration. So in this configuration, I have two cameras the projection center is C and C dash. Here this is the left camera and this is the right camera and this is the point the scene point is P x, y, z the scene point is x, y, z.

And distance between the two camera projection center that is the baseline the baseline I am considering the distance is B and if you see the image plane here this is the image plane corresponding to the left camera and this is the image plane corresponding to the right camera. So, what will be the size of the image in the left image the left image is u and in the right image it is u dash and I have considered the focal length the focal length is f.

So, from this configuration I want to find the depth information. So this process is called the triangulation. So, triangulation is a process of obtaining the 3D dimensional scene points from its projected points in the two image planes. So, I want to get the 3D information of the scene point from its projected points in the two image planes. So, if I have two image planes so from the two images I want to get the 3 dimensional scene point that means I want to determine the x coordinate, y coordinate and the z coordinate.

(Refer Slide Time: 25:25)



So, same configuration I am showing here. Now, I am defining a term that term is called the disparity. Disparity in the horizontal direction between the points u and u dash that is u minus u dash. So, in this diagram you can see similar right angle triangle. One is this triangle that is the UCB that is one triangle and another triangle is CPA that I am considering the CPA two right angle triangle and corresponding to the point the coordinate is x, y, z.

So, suppose this side is positive x and the left side is negative f corresponding to this point. So, if I consider this two triangles one is UCB and another one is CPA from similar right angle triangle I can write like this u by f is equal to minus h plus x divided by z. Why it is? So from this to this point from this point to this point the distance is h and from this what is the distance? This is x you can see this distance is x because this coordinate this is suppose the axes is 0 here, but in this case I am considering the negative direction.

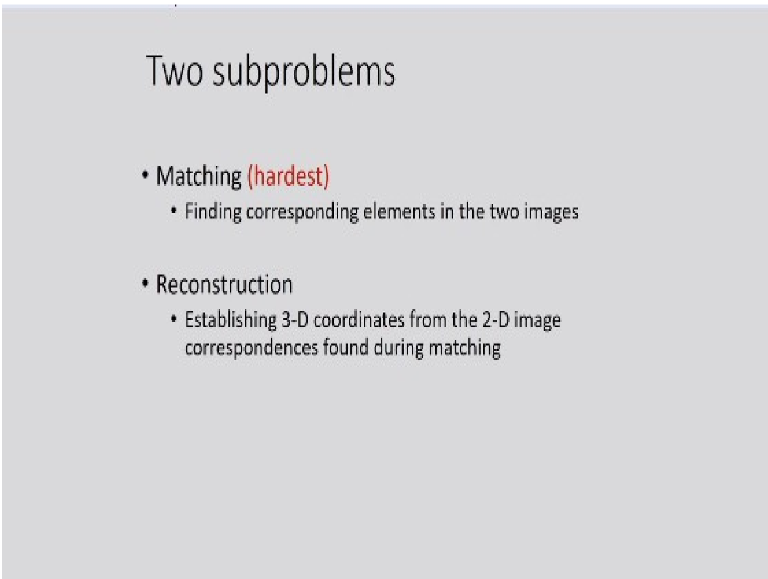
The negative direction is minus I have shown so that is why I have considered a negative sign here. So corresponding to this the distance from this point to this point if you consider this point to this point the distance will be h plus x, but that is in the negative directions so it is minus h plus x divided by z. Similarly, if I consider another triangle the triangle is this triangle, this triangle and also this triangle.

Corresponding to this what I am getting u dash divided by f, f is the focal length and I have to find the distance this distance I have to find that is in the positive direction. So it is h minus x that distance will be h minus x in the positive direction so this is the case. So, I am getting this two equations from this two right angle triangles and in this case from this two equations I can determine the z information, z is the depth information.

Z is equal to bf divided by d that I can determine. What is b? B is the distance between two camera projection centers. What is f? F is the focal length of the camera and what is d? D is the disparity in the horizontal distance that is u minus u dash. So, you can see that is the depth information. So, if the disparity is more the depth will be less. So, suppose the object suppose one object that is very close to the camera then in this case the disparity will be more.

And corresponding to this the depth will be less. So that means here you see the disparity is nothing, but the horizontal displacement of the matching points in the disparity value for this point. So, after determining the depth you can determine the rest of the coordinates, the rest of the coordinates is x coordinate and the y coordinate. So you can see how to define the depth information in this stereo vision setup.

(Refer Slide Time: 29:01)



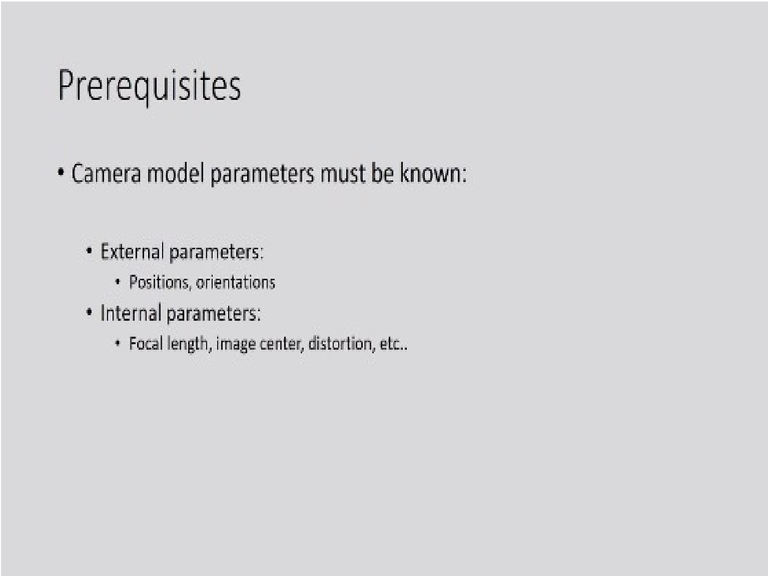
Now in this case if I have to find the disparity between the two images. So up till now I discussed about the stereo vision setup and I have discussed the concept of the epipolar planes the epipolar lines and also the concept of the general camera configurations. So after image rectification I am getting the canonical camera configurations. So, in the canonical camera configurations the epipolar lines will be parallel.

And that is after rectification the 2D search problem is converted into 1D search problem. After this I discussed about how to get the depth information from the stereo vision setup. Now after this what we have to consider, we have to determine the disparity we have to determine from the two images. One is the left image another one is the right image. For determining the disparity I have to find the matching point.

One is the left image another one is the right image so I have to determine the matching point. So, that is why the first I have to find the matching. After doing the matching, I have to go for the reconstruction of the 3D information. So, from the two images I can find the disparity from pixel-to-pixel from all the pixels I have to determine the disparity. So, if I find the disparity for all the pixels of the images then in this case I will be getting the disparity map.

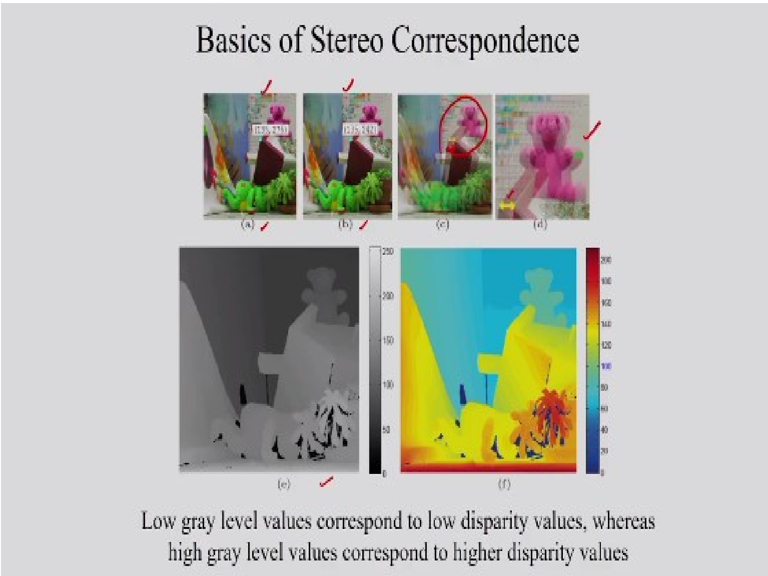
And from the disparity map, I can determine the depth map the depth information I can determine that is the reconstruction. So, here I have shown one is the matching another one is the reconstruction. So, finding the correspondence between the two images.

(Refer Slide Time: 30:35)



And in case of the matching this I have to consider that camera model parameters must be known. So already I have discussed about the camera parameters the extrinsic parameters of the camera external parameters of the camera that is the position and the orientation of the camera with respect to the world coordinates and internal parameters of the camera like the focal length, image centers and the lens distortions. So I have to consider this parameters are known.

(Refer Slide Time: 31:05)



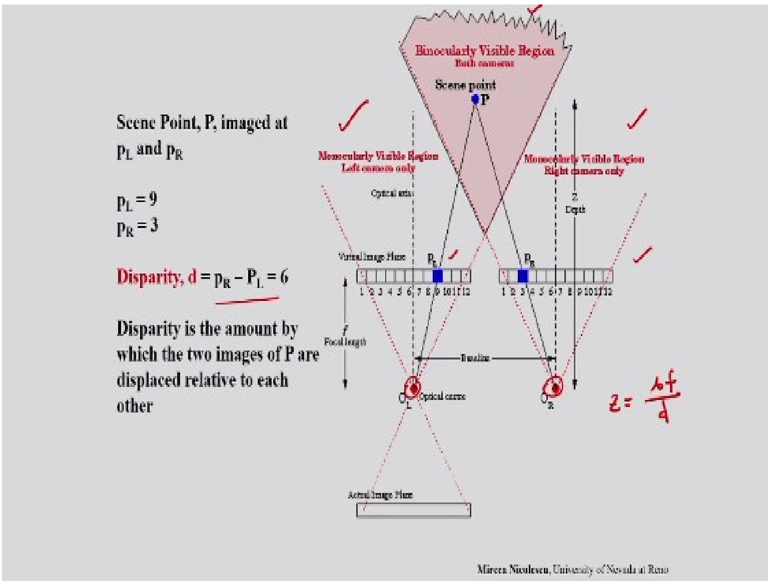
So, here you can see I have considered two images one is the left image another one is the right image and I want to show the disparity what is the disparity? In this case you can see in the diagram c I have shown the overlapping between these two images. I am just doing the overlapping of this and you can see the disparity in the two images. This yellow lines I have shown the disparity and this portion I have magnified in this image.

So, you can see the disparity here the yellow this arrow it shows the disparity between the two images the left image is shown in a and the right image is shown in b. Now from this you can see the disparity I have shown here in figure e. So disparity in terms of gray level intensity value. So, low gray level values correspond to low disparity values whereas the high gray level values corresponds to higher disparity values.

So, one point I have explained that is the concept of the disparity objects nearer to the camera encounter more shift compared to the more distant object and this is reflected in the disparity values that means the objects nearer to the camera have high disparity values. While the objects further from the camera have low disparity values. So that means the object nearer to the camera the depth will be less and object further from the camera that is away from the camera have low disparity value and that means the high depth value.

So, you can see the relationship between the disparity and the depth the same thing I have shown here in the color map showing the disparity values in the color map.

(Refer Slide Time: 32:56)

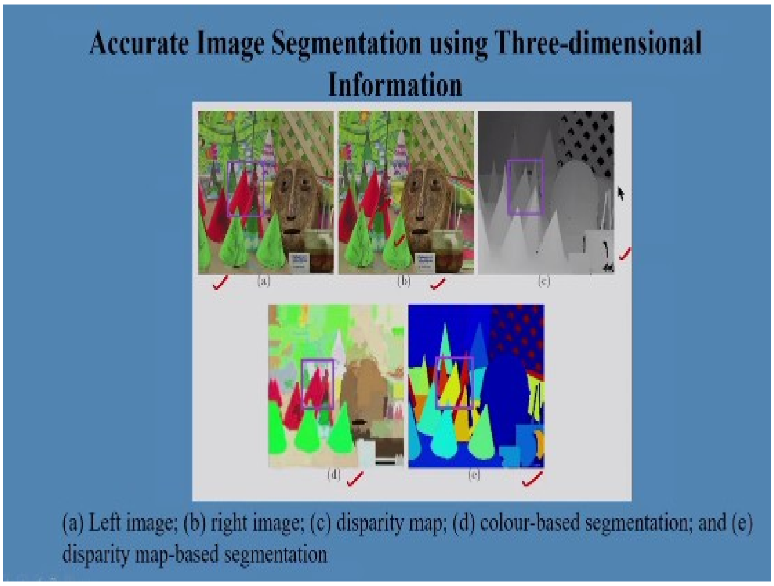


In this figure, I have shown again the same concept if I consider this portion, this portion is binocularly visible that means this portion is visible from both the cameras. I have two cameras one is the left camera, the left camera projection center is this, right camera projection center is this. So, this portion is binocularly visible, but if I consider this portion or this portion that is monocularly visible.

And I have the image plane the virtual image plane is here this is the left image plane and this is right image plane. Corresponding to the scene point P, what is the value in the left image and what is the value in the right image, it is 3. So, from this you can find the horizontal displacement, the horizontal displacement is pR minus PL that is 6 so that is the disparity is 6.

So, from this disparity we can determine the depth information by using that equation. So already the equation already I have explained so what is the equation z is equal to bf divided by d, d is the disparity and b is the distance between the camera projection centers. So, by using this equation you can find the depth information.

(Refer Slide Time: 34:09)

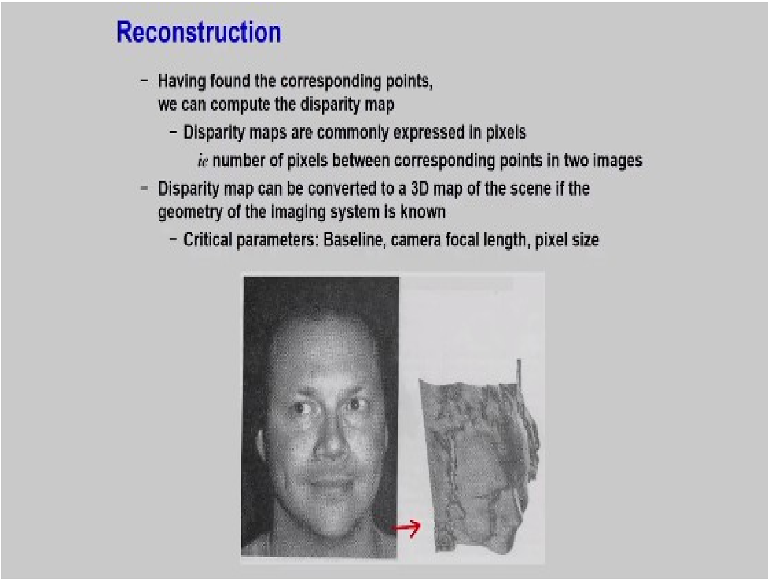


And this disparity map is quite useful for some applications like one application I am showing here. The image segmentation by considering the depth so here I have shown this example the image segmentation by considering the 3D information that means by using the disparity information how to do the segmentation. So image segmentation is nothing, but the partitioning of an image into connected homogeneous region.

So, in this case I have considered two images, one is the left image another one is the right image and in this case I have shown c is the disparity map I am obtaining from this two images I am getting the disparity map and this is the result of the color based image segmentation and this is the result for e is the result for disparity map based segmentation. So, from the disparity we can do the segmentation because the depth of this objects are different.

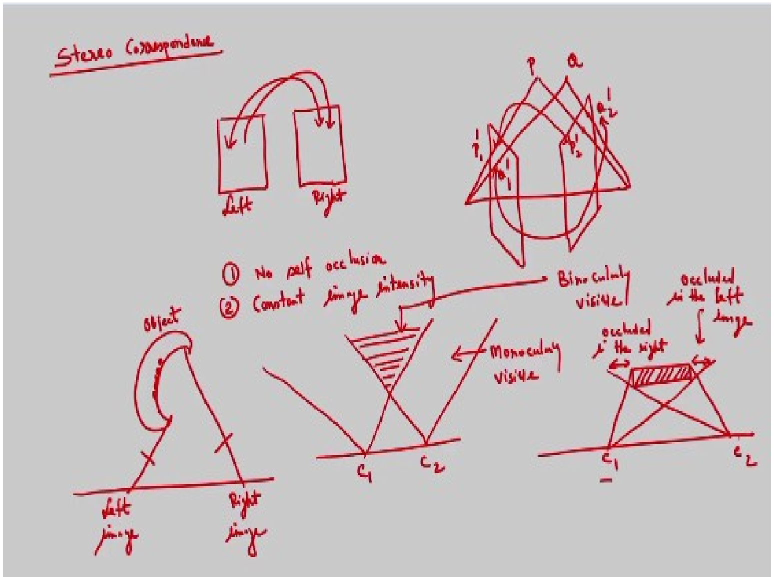
So, if I consider depth of this and depth of this depth will be different and based on this disparity information we are doing the segmentation. So, accurate segmentation using three dimensional information.

(Refer Slide Time: 35:16)



And here I have shown after determining the disparity maps we can find the 3D map that is the disparity map can be converted into 3D map and if we know the geometry of the imaging system that means the parameters like the baseline, the camera focal length, pixel side if you know then this I can get the 3D information from the disparity map, so that is called the reconstruction.

(Refer Slide Time: 35:43)



Now for stereo correspondence so already I told you that for a stereo correspondence I have to find the correspondence between the two images one is the left image another one is the right image. So, I have to find the correspondence between the points like this. One is the left image and one is the right image. This is called the stereo correspondence and in case of the stereo vision setup I can show you that same thing I am showing here.

So, this is my left image and this is my right image. So, I have two points suppose P points and the Q points. So this point is P suppose and this point is Q so corresponding to the first P point I have the P1 dash and corresponding to second point I have the projection that is Q1 dash and also corresponding to this left, right image I have P2 dash and I have another points that is the projection is Q2 dash.

So, in this case I have to find the correspondence between this points the point is corresponding between this points P1 and P2 I have to find the correspondence. Similarly, I have to find the correspondence between the points this point I have to find the correspondence. So, in this case for finding the correspondence one important point is there should not be no self-occlusion.

If I want to find the correspondence there should not be any self occlusion and also this point is important constant image intensity suppose the constant image intensity is there then it is very difficult to find the correspondence between the points. So, if I consider a flat object or maybe the non textured object or maybe the white object or maybe the uniform brightness if I consider then in this case it is very difficult to find the correspondence between the points.

So, this self-occlusion I can give one example is something like this. Suppose, if I consider object something like this one object. So, here this is the left image I have the left image and this is the right image I am getting. To see this portion is not visible from the cameras if you see that is one example of the self occlusion.

So in this case it is very difficult to find the correspondence between the images. This is one example of the self occlusion. So, I am considering this is the object and in this case these points are not visible by the camera the left camera it is not visible, but right camera it may see this points. So, in this case it is very difficult to find the correspondence between the images.

So that means I can show in this diagram so already I have shown so it is C1 is one camera from projection center another one is C2. So this portion if we consider this portion is binocularly visible and this portion is monocularly visible and this portion is binocularly visible. One is monocularly visible another one is the binocularly visible and I can show another example suppose so I have C1 camera another one camera is C2.

I have two cameras similarly I have another camera C2. So, if I consider this portion is occluded in the left image because that portion is not visible by the camera C1 and if I consider this portion this is my object. So this portion is not visible by the right camera. So, that means this is occluded in the right image. So, I have shown this how to find the stereo correspondence between the two images.

So one problem is the self occlusion, another problem is the constant image intensity. So, if I consider the flat object, if I consider a non textured object or maybe the white object the object with uniform brightness then in this case it is very difficult to find the correspondence between the images.

(Refer Slide Time: 42:27)

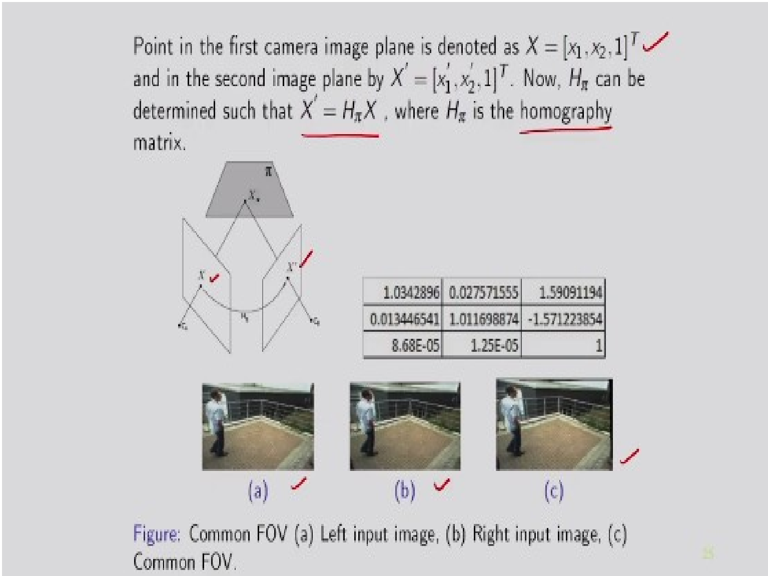


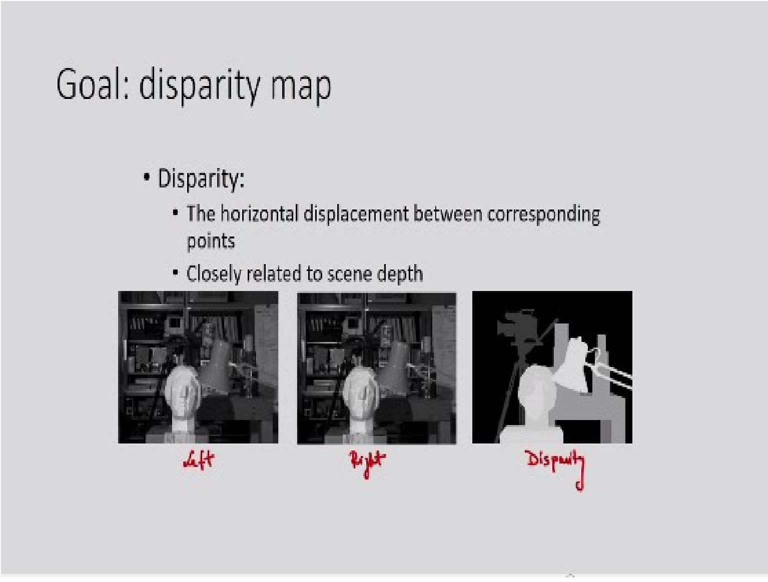
Figure: Common FOV (a) Left input image, (b) Right input image, (c) Common FOV.

So, in this case I have shown two images the first one is the left image another one is the right image. So point in the first camera image plane is denoted by X. So this is X the first image plane I am considering the point is represented by x1, x2, 1 that is in the homogeneous coordinate system and second point the second image plane is represented by x1 dash, x2 dash and 1.

So, I am considering the second image plane that is the point is represented by x dash so x1 dash, x2 dash, 1 that is in the homogenous coordinate system. Now this homography matrix H pi is the homography matrix. So, I can do some transformation so that x can be converted into x dash by using the homography matrix. So, here I have shown one example the left image this is the right image.

And in this case I have shown the common field of view that this portion is visible by both the cameras. So, I have shown the one example of the common field of view.

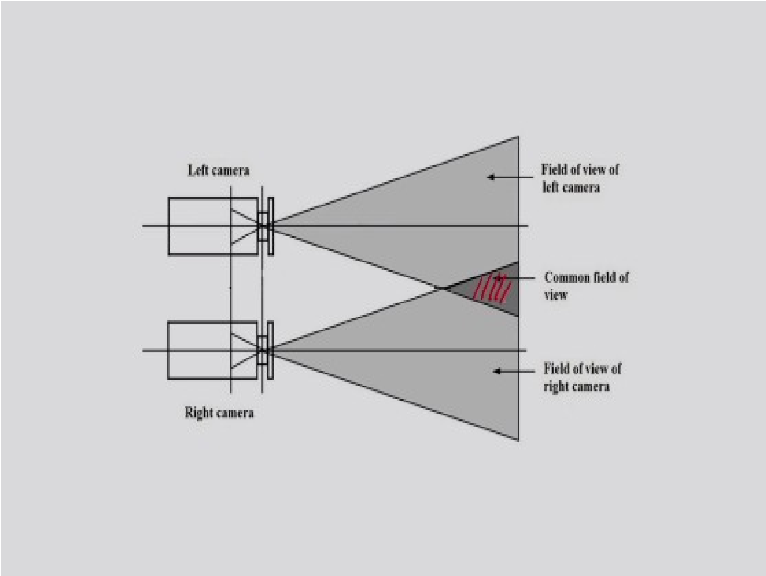
(Refer Slide Time: 43:39)



And already I told you the disparity map so disparity means the horizontal displacement between corresponding points. So, after determining the disparity we can determine the depth information. So, this is my left image, this is my right image and this is the disparity map. So, for all the pixels of the image I am determining the disparity then after determining the disparity that is the disparity map I can determine.

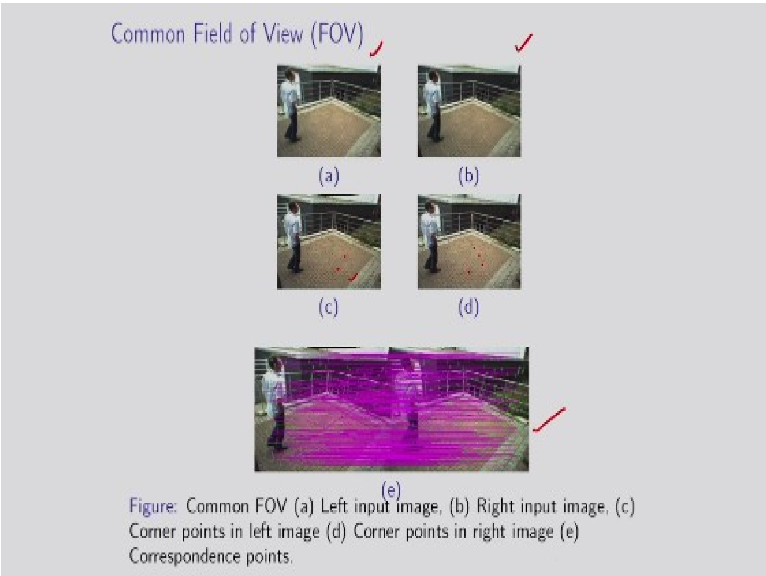
So for all the pixels of the image the left image and the right image I am finding the disparity. From this disparity I am getting the disparity map and from the disparity map I can do the 3D reconstruction.

(Refer Slide Time: 44:20)



So, in this diagram I have shown the common field of view the two cameras I am considering the left camera and the right camera and I have shown the portion that is the common field of view I have shown.

(Refer Slide Time: 44:30)

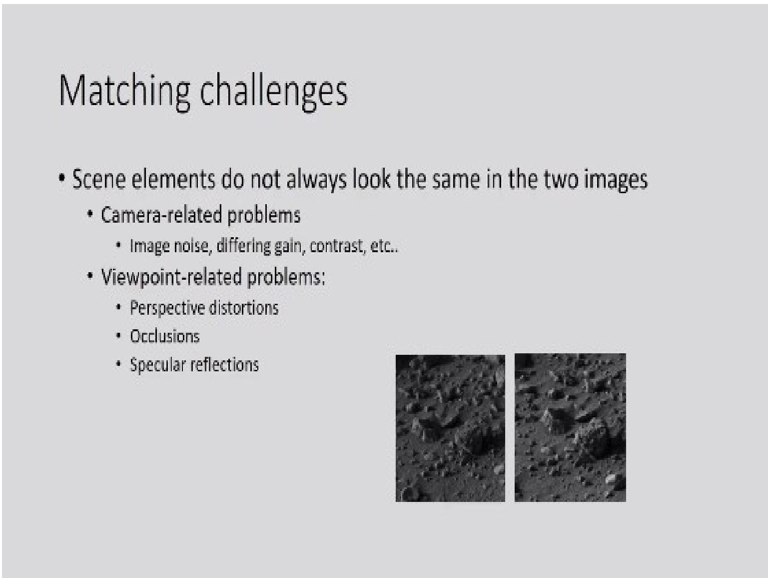


Here I have shown again the same example the left image this is the left image next one is the right input image and after this I am showing some corresponding points the corner points I have shown some red points you can see, red points here I have shown some red points that is the corner points and after this I want to find the correspondence between this points. So, in this case I am finding the correspondence between this two points.

That is corresponding to the common field of view I am finding the correspondence between the points of the left image and the second image. Now, let us consider the matching problem. So, in the matching problem, I will consider two approaches one is the pixel based approach and another one is the feature based approach. So, if I want to find the correspondence between the images the left image and the right image.

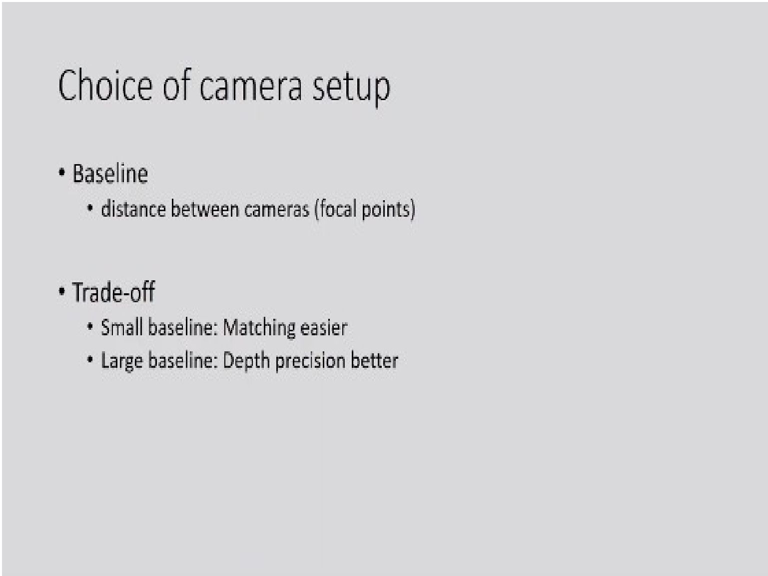
So, I can compare the pixel intensity value and based on the pixel intensity value I can find the correspondence that is called the pixel based method. Another method is the feature based method. In feature based method, I can extract some features the important features in the images and based on this features I can do the matching. So, this two approaches I am going to explain.

(Refer Slide Time: 45:44)



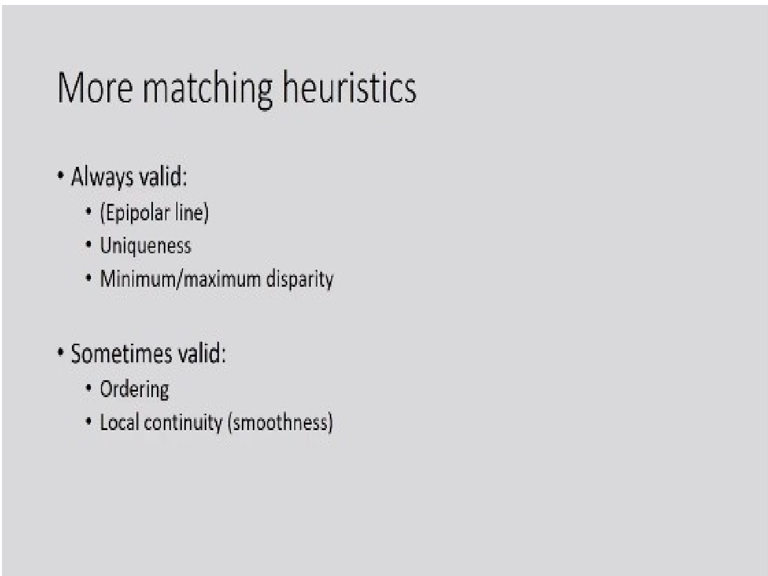
Now there are many challenges for matching. Some challenges like image noise differing gain of the cameras, the contrast and also some constraints like the perspective distortions, occlusion, specular reflections. So this challenges I am going to discuss one by one.

(Refer Slide Time: 46:04)



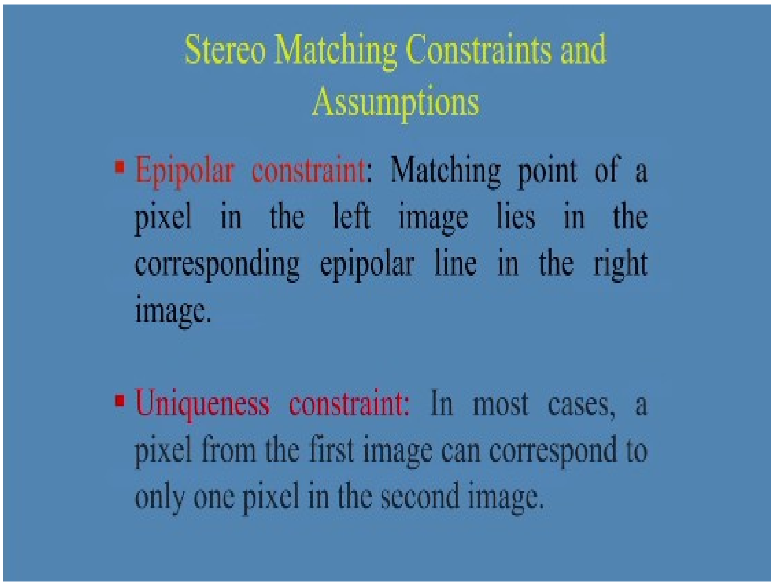
And in this case the choice of camera setup is very important. The baseline is the distance between the camera, the camera means the center of the projections of the camera the left camera and the right camera and if I consider small baseline then in this case the matching will be easier and if I consider large baseline then depth precision will be better.

(Refer Slide Time: 46:27)



And also I have to consider some assumptions for stereo matching. These assumptions are like epipolar constraints, uniqueness constraints, minimum and maximum disparity, ordering constraints, local continuity constraints. So, there are some constraints that means some assumptions I have to consider for stereo matching, So, I will explain all these assumptions, all these constraints one by one.

(Refer Slide Time: 46:52)

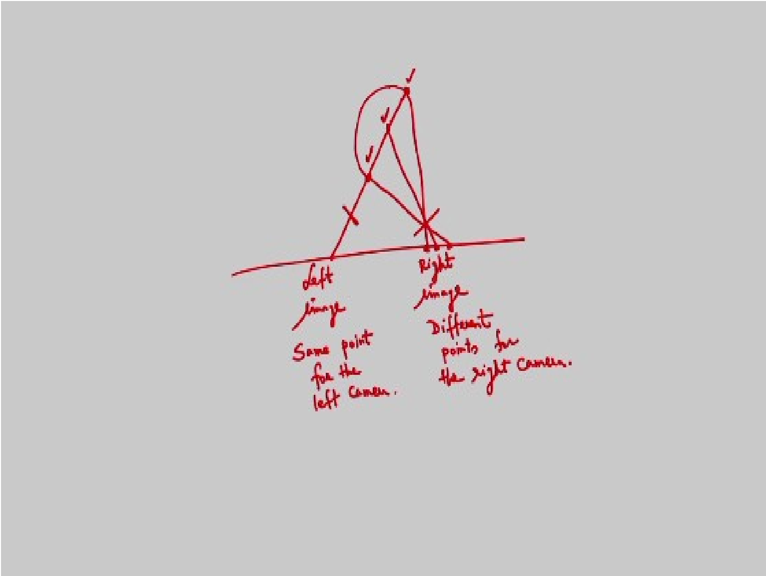


So, first assumption the first constraint is epipolar constraints. So, in this constraint the matching point of a pixel in the left image lies in the corresponding epipolar line in the right image that means already I have explained that is the 2D search problem is converted into 1D search problem. So in this case I have to find the correspondence along the epipolar line only. I have to see the correspondence along the epipolar lines.

This is called the epipolar constraints. Next one is I am considering the uniqueness constraint. In most of the cases a pixel from the first image can corresponds to only one pixel in the second image. This is most of the cases it is true, but suppose if I consider the opaque objects. The opaque object satisfy this constraints, but if I consider transparent object the transparent objects violate this condition.

And this is because of the fact that the many points in a three dimensional space is projected on to the same point in an image plane because of this condition is not satisfied. So, I am giving another example in which the uniqueness constraint is not satisfied.

(Refer Slide Time: 48:09)

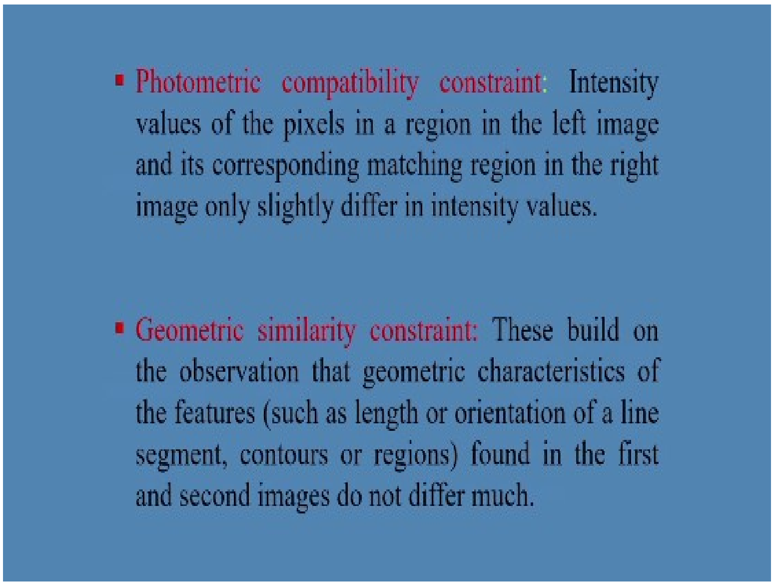


So, I am showing one diagram here. Suppose, I have one object something like this the surface and I have the left image and the right image. So, corresponding to this point I have this projection. So, this is my left image and suppose this is my line corresponding to this point I have this one, corresponding to this I have this one, corresponding to this I have the projections.

This is my right camera, right image. This is same point for the left camera, but in the right camera it is different points for the right camera. So, if I see here this point, this point, this point and this point it is same point for the left camera because I have the projection here. So, one projection I am getting, but for the right camera different points this point, this point and this point it will be the different points for the right camera.

So, in this case it is very difficult to find the correspondence between the left image and the right image and in this case is very similar to the case already I have explained the case is self occlusion. So, this is called the uniqueness constraint. So, as per the uniqueness constraint in most of the cases there exist at most one matching pixel in the right image corresponding to each pixel in the left image that is the uniqueness constraint.

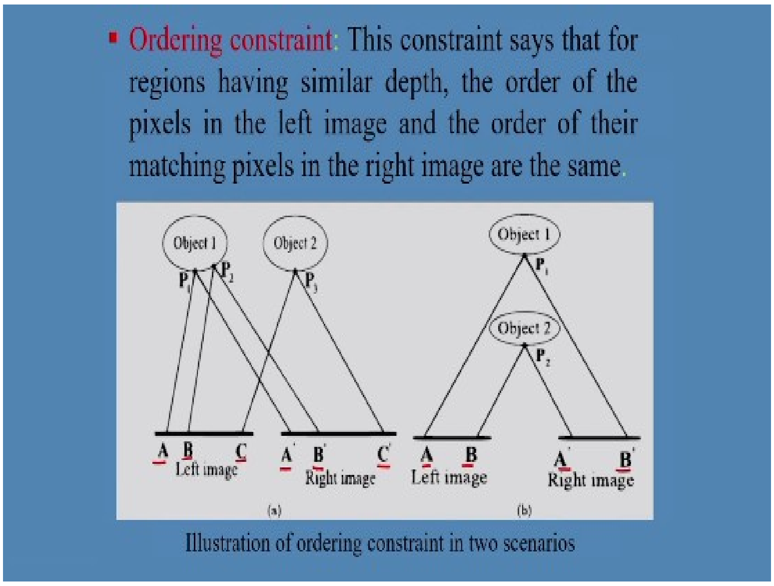
(Refer Slide Time: 50:30)



The next constraint is the photometric compatibility constraint. Intensity values of the pixels in a region in the left image and its corresponding matching region in the right image only slightly differ in intensity value. This slight difference in the intensity value is due to the different camera positions from where the images are captured. This is obvious and this is called the photometric compatibility constraints.

Another one is the geometric similarity constraints. So, as per this constraint if you see this observation is mainly that geometric characteristics of the features, the features maybe something like the length, orientation of a particular line, contours regions found in the first and the second images do not differ much. So, features which are available in the first image and with the features which are available in the second image they do not differ much that is the geometric similarity constraint.

(Refer Slide Time: 51:27)

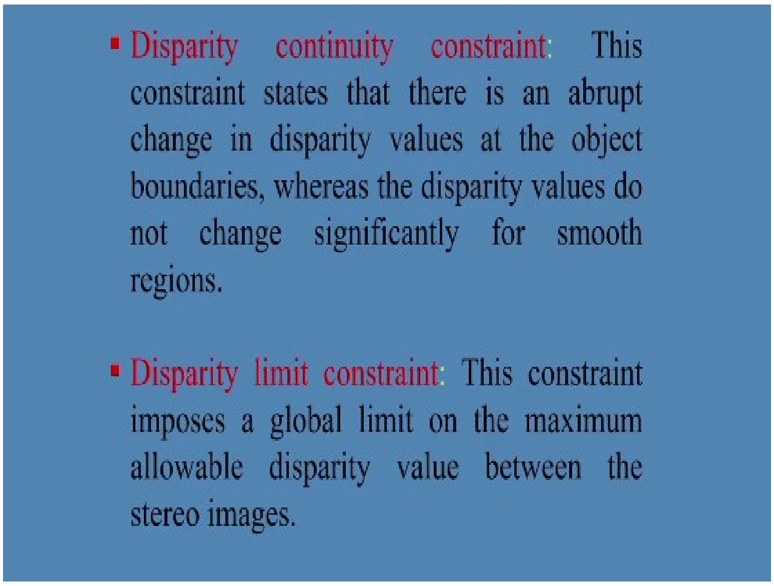


And this is very important the ordering constraints. This constraint says that for regions having similar depth the order of the pixels in the left image and the order of their matching pixels in the right image are the same. So, here I have shown two cases is A and B illustration of ordering constraint in two scenarios. So, in the first example the ordering constraint is fulfill. Why it is fulfill?

As the points A, B, C and their corresponding matching points A dash B dash and the C dash follow the same special order. So that means here I am considering A, B, C that is in the left image and A dash, B dash, C dash in the right image they follow the same special order that is the ordering constraint, but in the second case if you see this constraint feels at the order of the points A and B in the left image is different from the order of the corresponding matching points A dash and the B dash.

If you see this example the ordering constraint is not satisfied the second case. So, this is the illustration of ordering constraint in two examples the two scenarios.

(Refer Slide Time: 52:44)

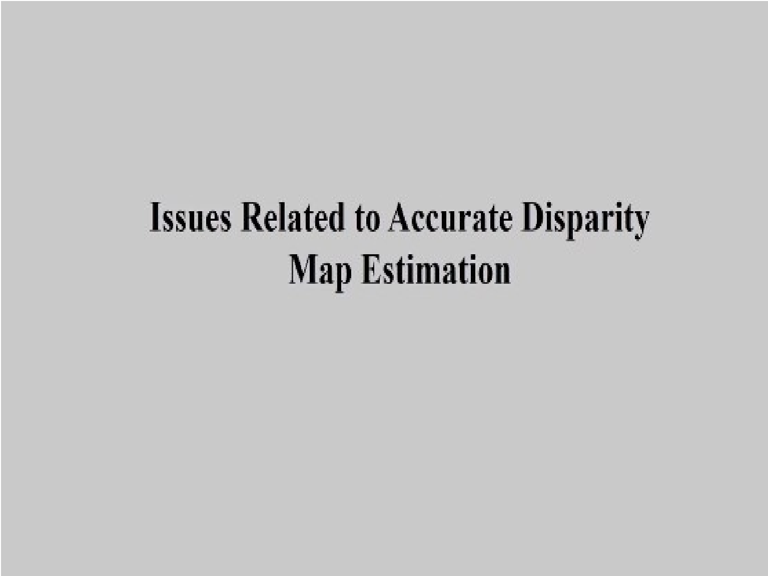


The next one is the disparity continuity constraint. So as per this constraint that there is an abrupt change in disparity value at the object boundaries that means if I consider the edges and the boundary there is an abrupt change in the disparity values whereas the disparity values do not change significantly for the smooth regions and this is quite obvious. For the boundaries or for the edges the disparity values changes abruptly.

But if I consider the smooth region the disparity values do not change significantly. This is called the disparity continuity constraint. Another one is the disparity limit constraint. This constraint imposes a global limit on the maximum allowable disparity values between the stereo images, but this is mainly based on psycho-visual experiments which says that the human visual system can only fuse the stereo images if the disparity values do not exceed a particular limit.

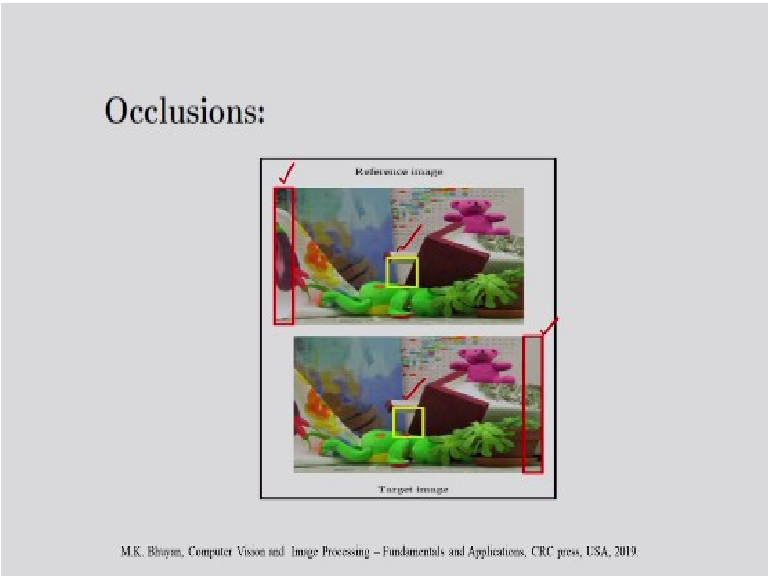
This is mainly based on psycho-visual experiments this is called the disparity limit constraints. So this constraints that this assumptions are quite important because based on these assumptions we are going to find the correspondence between the images.

(Refer Slide Time: 54:05)



The next one is the issues related to accurate disparity estimation. So, we will discuss some issues related to accurate disparity estimations.

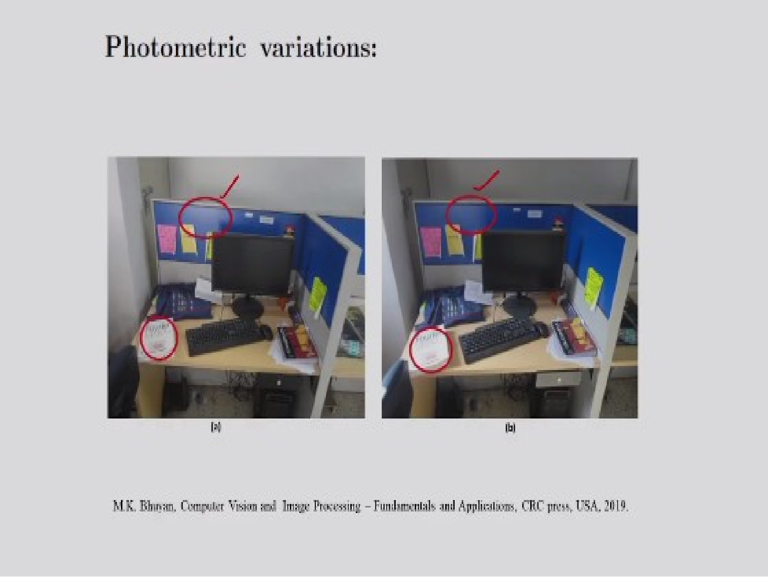
(Refer Slide Time: 54:15)



And one important issue is occlusions. Here I have given one example of the occlusions. Now occlusion is mainly because of different field of view of the cameras. So, in this example I have shown that this portion is only visible in the left image and that is not visible in the right image, but again if you see this portion is visible in the right image, but that is not visible in the reference image.

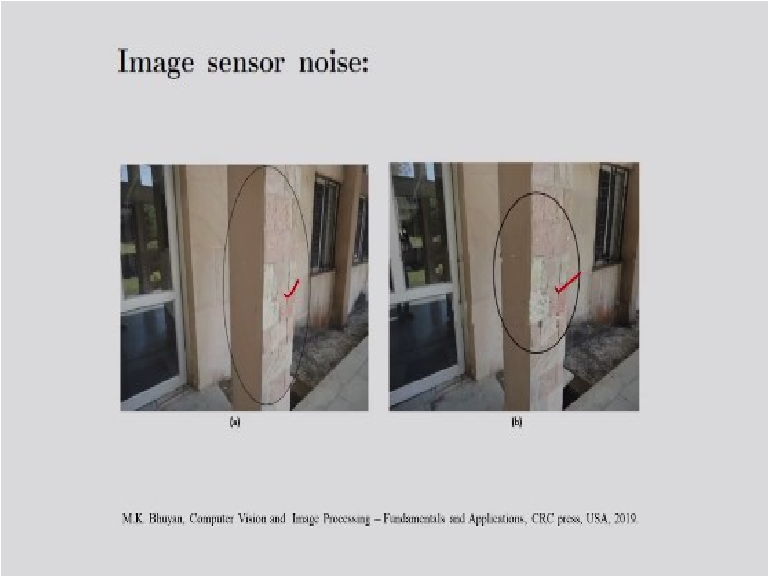
So then in this case it is very difficult to find the correspondence between this portion of the image. Also another case is that is the presence of overlapping of different objects that means if I consider overlapping of different objects located at different distances from the camera that is one case then in this case if I consider this portion the overlapping of the objects then in this case also it is very difficult to find the corresponding to that portion of the image because that portion is occluded.

(Refer Slide Time: 55:17)



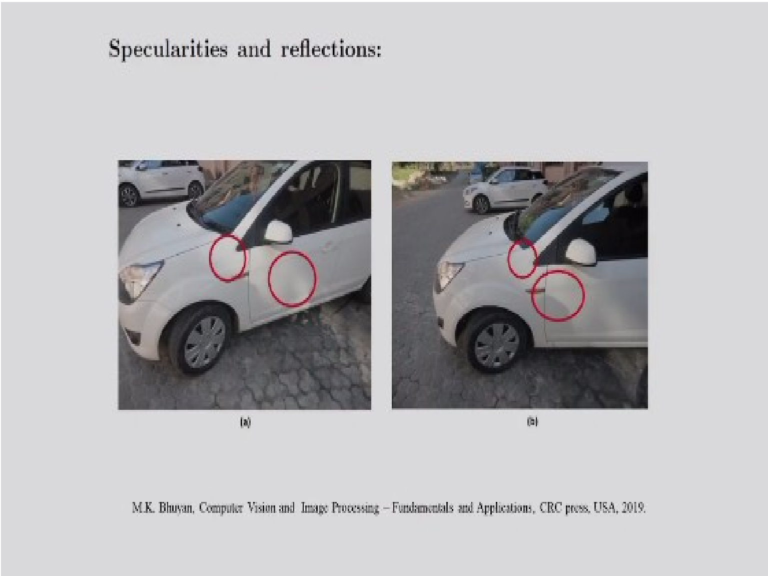
The next one is the photometric variations. Now the optical characteristics of the cameras may slightly differ and this leads to photometric variations of the stereo image pairs. So, in this case I have given this example you can see the photometric variations between the left image and the right image. Also, if I consider this portion you can see the photometric variations because of the optical characteristics of both the cameras may slightly differ.

(Refer Slide Time: 55:45)



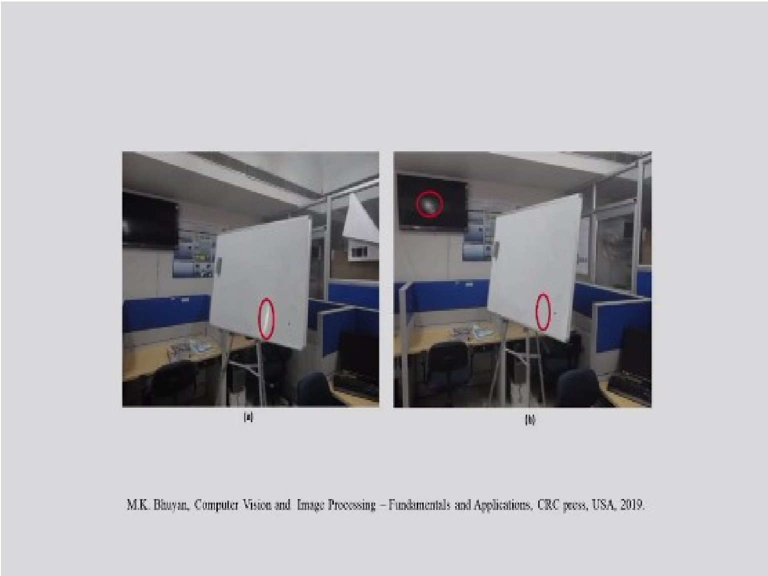
Now let us consider another problem that is the image senor noise because of the noises in the camera you can see this noise is here and noise present in the image then also it is very difficult to find the correspondence between the images. So, one technique we can apply that is the preprocessing of the image that is you can remove noises in the image by filtering.

(Refer Slide Time: 56:06)



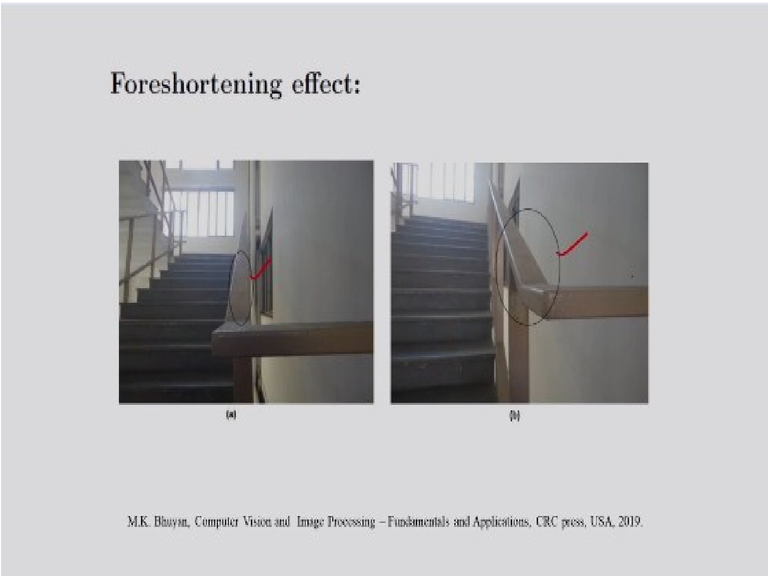
The next point is the specularity and the reflections. So, already I have explained what is the specular reflections? For a specular surface the radiance leaving the surface are dependent on angles. So the disparity map obtained from the stereo image pairs may not give actual information of the specular surfaces. In this example I have given the specular surfaces, this is the specular spaces, this is the specular surface like this I have the specular surface then in this case also very difficult to find the correspondence between the images.

(Refer Slide Time: 56:37)



I am giving another example here you can see the specular surface. The specular surface is nothing, but the mirror like surface. So, mirror like surface is there and in this case also it is very difficult to find the correspondence between the images.

(Refer Slide Time: 56:51)



One important effect is the foreshortening effect. I explained the concept of the foreshortening factor. The foreshortening factor is cos alpha or cos theta. As per effect what is the effect? The appearance of an objects depends on the direction of the view point that concept I have explained already. Hence, an object may appear compress and occupies smaller area in one image as compared to the other image.

So, in this example I have shown. Here you can see the area is less, but you can see the area is more corresponding to that portion of the image then in this case it is very difficult to find the correspondence between the images.

(Refer Slide Time: 57:34)

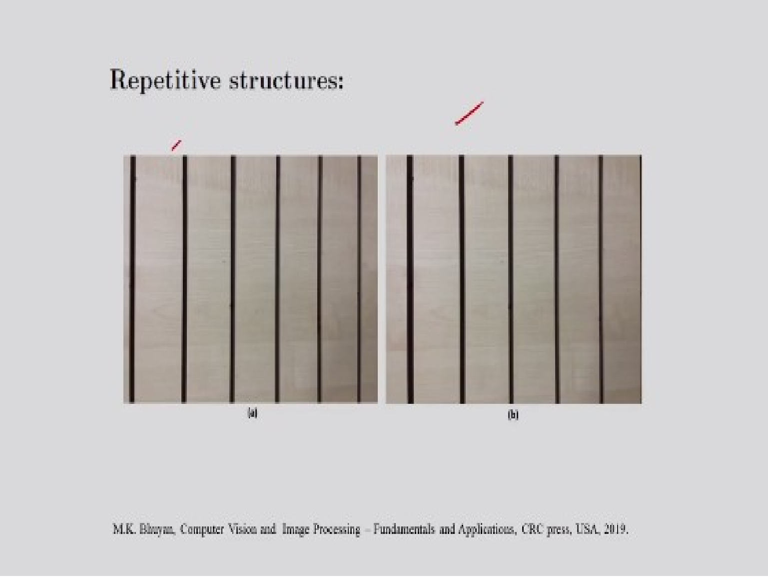
The next problem is the perspective distortion. The perspective distortion is a geometric deformation of an object and its surrounding area. This is mainly because of the projection of a three dimensional scene onto a two dimensional image plane that is nothing, but the 3D to 2D projections.

So because of this perspective transformation the perspective transformation makes an object to appear large or small as compared to its original size. If you see in this image this portion looks small, but if you see this portion, this portion looks large because of the perspective distortions.

(Refer Slide Time: 58:13)

And also another problem the textureless regions, in this example I have shown the textureless regions, this is one region, another region then in this case it is very difficult to find the correspondence between the images corresponding to that portion of the images.

(Refer Slide Time: 58:28)



And here I am giving one example of the repetitive structures what is the repetitive structures you can see? This structure is repeated and in this case also it is very difficult to find the correspondence between the points of the left image and the right image.

(Refer Slide Time: 58:43)

And next point is the discontinuity. So, what is discontinuity? Generally we considered that assumption is that surface is present in the scene are smooth. This is the assumption we have considered. However, this constraint is not fulfilled when multiple objects are present in the scene. So, if I consider this here the multiple objects are present here because of this we have the discontinuities between the different objects in a scene.

And there is an abrupt scene in a disparity value in the boundary regions. So this is called the discontinuity. You can see the discontinuity is because of the multiple objects present in the scene. So in this example I have shown this portion is discontinuous I have shown the discontinuous region here this is one problem of finding the disparity map. So, up till now I discussed about the concept about the stereo vision.

So, first I discussed about the general camera configuration. In general camera configuration, epipolar lines are not parallel after this I discussed the concept of image rectification. In the image rectification, the epipolar lines will be parallel and in this case so 2D search problem is converted into 1D search problem. So, in the first case I have considered the general camera configuration and after this I considered the canonical camera configuration.

Now because of this configuration the canonical configuration the epipolar lines will be parallel, then the 2D search problem is converted into 1D search problem. After this, I discussed about the concept of the disparity map, so how to find the disparity maps and from the disparity maps I can determine the depth information. So depth point is very important.

So, for all the pixels of the image the left image and the right image I can find the disparity values.

From the disparity values, I can find the disparity maps and from the disparity map I can determine the depth map. After this, I discussed about some assumptions to find the correspondence between the images the left image and the right image like epipolar constraint, discontinuity constraint. So there are many constraints we have discussed and after this problem of finding the matching points for the stereo correspondence I discussed some problems like the problem of occlusion, perspective distortions.

So these problems are quite important. So, in the next class I am going to discuss the concept of stereo matching. One approach is the pixel based approach and another approach is the feature based approach. So, next class I will discuss these concepts. So, let me stop here today. Thank you.