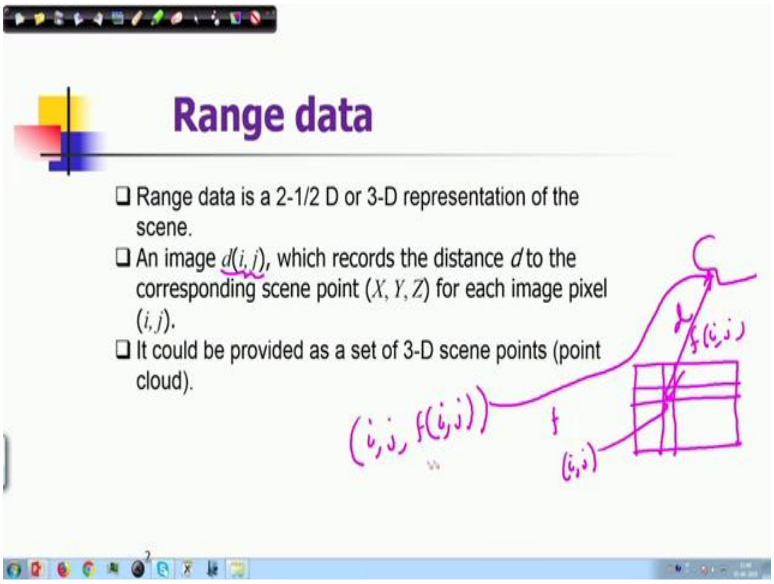
Computer Vision

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

Lecture – 41
Range Image Processing – Part I

In this lecture we will start a new topic to discuss; this is on Range Image Processing.

(Refer Slide Time: 00:23)

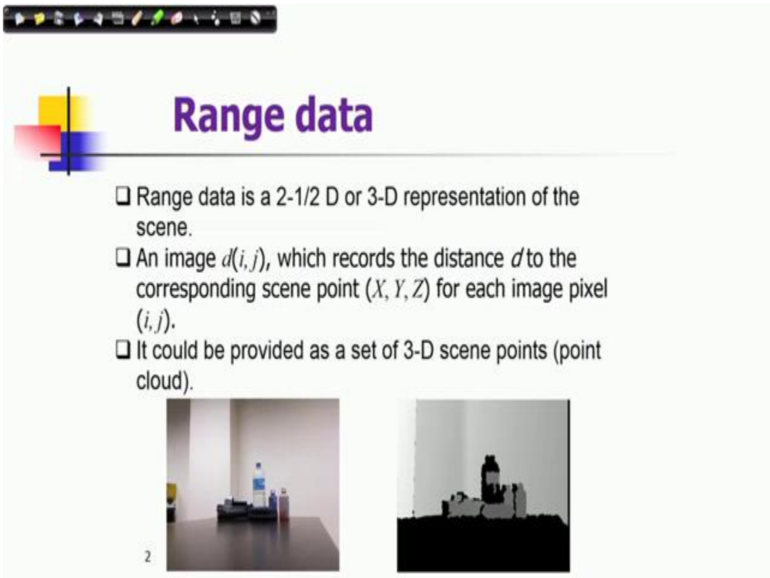


So, let us first understand what is meant a range image. A rang data is a $2 - 1/D$ two and half D or three-dimensional representation of the scene. It is called two and half D because you have only the surface information in the form of a data it is discretized pixels discretized points on the surface that is represented as an image. So, the representation is something like this that say consider this function say d(i, j) and it records the distance of the corresponding scene point. So, in an image when we consider the array representation of an image at pixel the value what we get in the range image is not the intensity, but the distance of the surface point.

So, this value d that is recorded at this point and this distribution; this distribution of this surface points over this discretized space that is a functional distribution that give me the range data, or we sometimes we also consider this data as depth data that is also another nomenclature we used for range data. Also it could be provided as a set of 3-D scene points or point cloud.

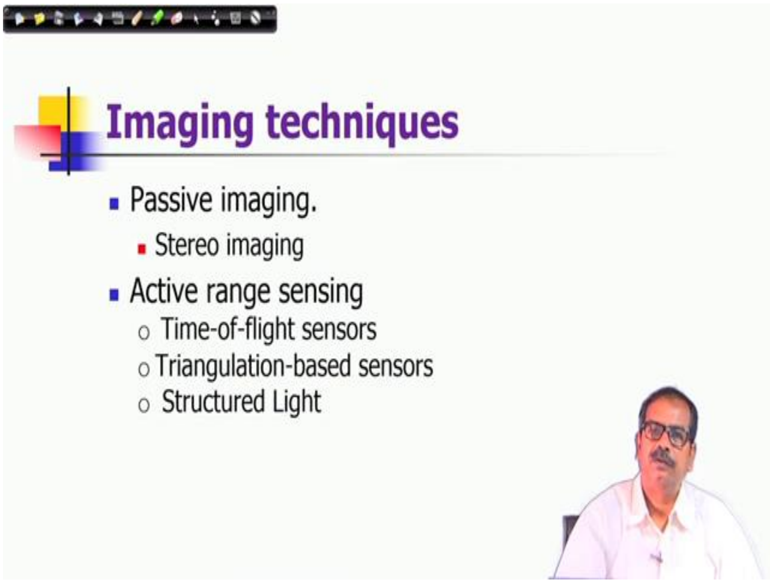
So, just to explain once again that suppose this position is array index position is given (i,j) and the corresponding functional value here it is f(i, j), then the x, y, z coordinate of the pixel value is given as(i,j), f(i, j) in a discrete 3 dimensional space so this is how this corresponding location is given. So, if I consider a collection of all these points in this representation in this form then we get a set of 3-D scene points.

(Refer Slide Time: 02:18)



We will have range data is shown here in the left side corresponding intensity image is shown and the right side its data range data captured using the Microsoft Kinect sensor that has been shown here.

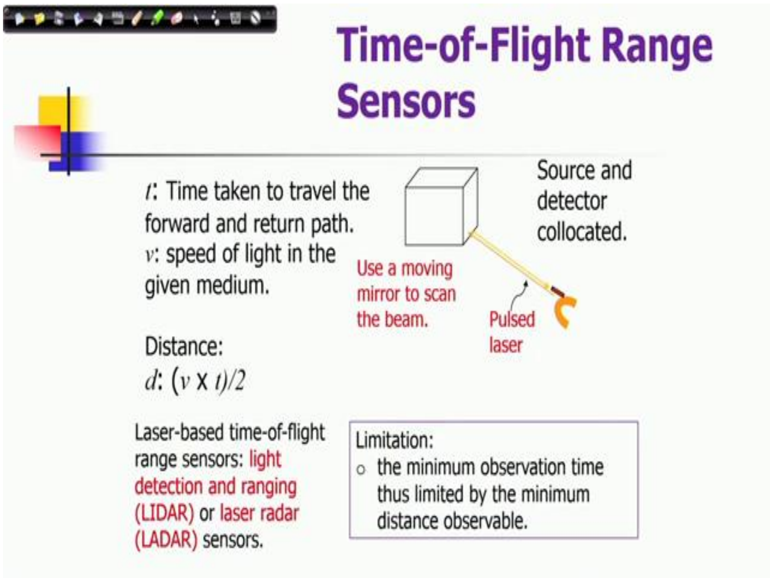
(Refer Slide Time: 02:35)



There are various principles on which range imaging takes place we have already discussed about stereo imaging system and we discussed how 3 dimensional scene information could be computed using a stereo imaging system. So, this kind of system is passive imaging system and which would provide you also the range image, but there are other kinds of systems where you get high resolution range image and those are active range sensors or active range sensing mechanisms that is followed.

Active in the sense that here you need to use an energy source apart from the scene for detecting the for measuring the depth you have to project a ray from a particular source of energy and by illumination you have to consider the computation of the depth. So, there are three different types of active sensing mechanisms are there; one is time of flight sensors the other one is triangulation based sensors and structured light based sensors. So, let me explain those principles one by one.

(Refer Slide Time: 03:50)



So, in a time of flight range sensors there is a light source and it say a laser light source usually and you project a light or you transmit a pulsed laser from that point and again he reflection gives you back the corresponding laser signal. So, by detecting that pulses you can measure the time of flight or the duration of the transmission and reflection of the signal. So, if you know the velocity of light , all of we know the velocity of light, then you simply you multiply with that time that would give you the twice of the distances.

So, that is a principle; you have both source and detector co-located because the point from where you are transmitting you are also detecting at the same location. So, directly you are getting the shortest path from that location since light travels through the shortest path in that direction itself, you can get this depth. So, in this relationship is very simple if you can measure this time, then you multiply with speed of light and then you get the twice of the distance we have to divide by two to get the actual distance.

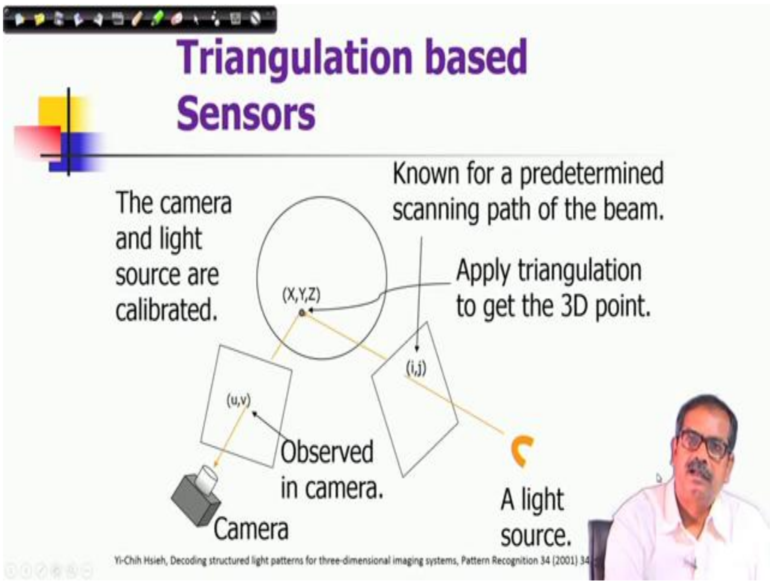
So, this light source as I, as it is shown here pulse laser. So, it is a laser based time of flight light it is a range sensors there are different kinds of lasers are there like light detection and ranging red lasers are there sensors are or laser ladar sensors are there and these are quite popular lidar or ladar based sensors. And since you are just measuring only one surface point and the depth of a one surface point at a given direction.

So, you have to scan the you know whole surface by maneuvering about the direction of the laser beam and there are moving mirror, you can use moving mirror to scan the beams. So, there are mechanisms by which the mirror angles could be varying and accordingly the direction of laser beam could be varied and in a given predetermined path you can find out that what is the depth from where the reflection came and that is how you can get the surface points.

So, there is a limitation of this kind of sensing you are limited by the minimum observation time because the mechanism by which you know you can detect a pulse laser, it has its width laser pulse duration, duration of the pulse that is one thing and another thing is the sampling intervals of emitting pulse laser.

So, these are you know limiting; these are limiting your observation time and does know your minimum distance what is observable that is also limited by that minimum observation time.

(Refer Slide Time: 07:08)



The other kind of sensor; other kind of principle is triangulation based sensors where you can see in this particular diagram that there is a source for laser beam, it could be any light source, but usually lasers are used for in this case also. So, you project that light on a particular surface point and then the reflection of that light will be captured by a camera.

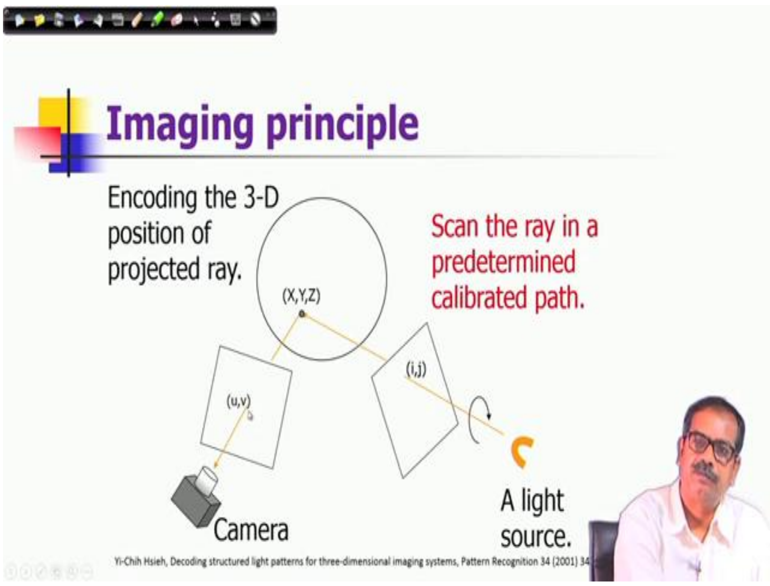
So, in the camera from the images you can get the point of the reflected path, if both camera and light source they are calibrated then you can get the equation of this two lines as we have discussed in the previous lectures of stereo geometry. That if I can get the directions of this 2 rays in 3 dimension then the intersection of these two lines will give you the corresponding three dimensional point.

So, this imaging system it calibrates everything. So, from the camera calibration we can find out that in which pixel of the camera in image plane which is illuminated from where the light reflected light is coming to the camera. So, you can get from the camera parameters the equation of this particular you know direction of ray in the equation of this ray in the three-dimensional coordinate system.

Similarly by knowing the coordinate location of the laser beam and also by knowing also the in the calibrated face by knowing also the value of i, j with the when you are you know when you are sending this transmitting this laser or transmitting this light that direction itself is also encoded. Or it is also known in the in the system then you can get the equation of these particular direction particular ray and by solving them you can get the three dimensional scene information.

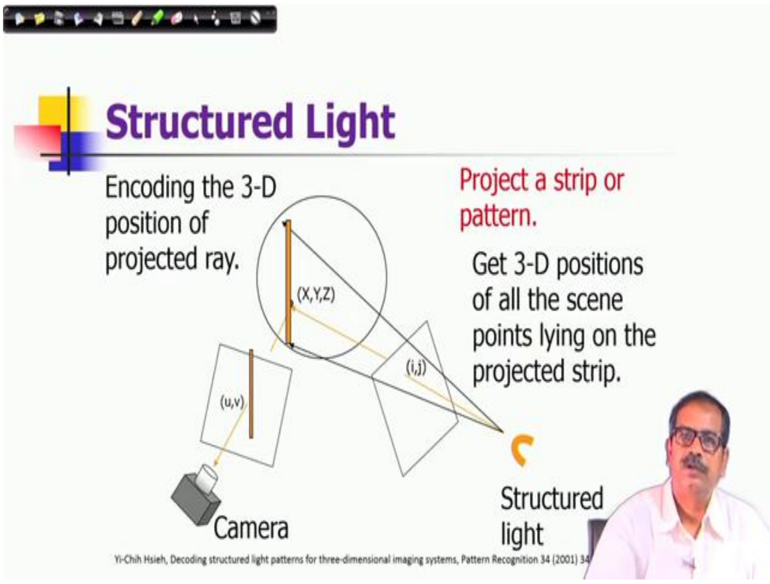
So, there are various ways by which you can determine what is the directions by which this particular along which this ray has been transmitted. So, those variations are there. So, this is what known for predetermined scanning path of the beam and then solving the equations and this is observed in camera and you apply triangulation to get that 3-D point.

(Refer Slide Time: 09:55)



Now, this same principle is used because as I mentioned only for one directions at a particular time you can get informations only for one surface point given a direction. Now, you have to vary these directions you have to scan the corresponding you know projected ray of the, from the light source projected from the illuminated in a predetermined path. Or in a calibrated path over a at every instance you should know the coordinate of i, j in image plane which will give you the corresponding equation of the straight line equation of this ray. And also you are observing the image points in the camera because of this illumination and for that illuminating point we will discriminate we will help you easily identify this particular point and you can apply this triangulation law.

(Refer Slide Time: 10:50)

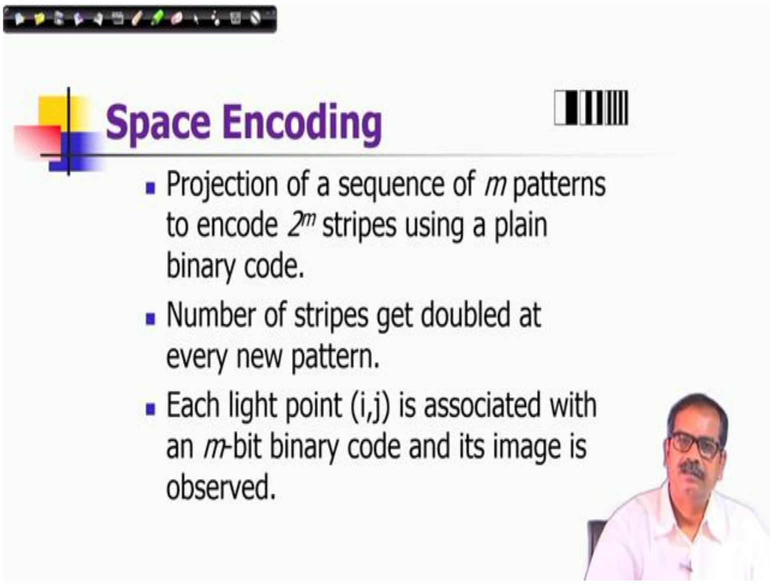


Now, the technique of you know triangulation using the principals what I discussed it is quite slow because you know you have to consider shooting ray for every point in its corresponding you know plane of in the respective image plane if I say. So, this is a if I say this is the plane where you are recording the you are identifying a particular directions like it acts like an image plane of a camera in the case of structured light or in the case of light source. So, instead of projecting single ray, what you can do? You can project a vertical you know light itself vertical strip on the surface.

So, which means that all your all the directions along this vertical line they are encoded into the strip and when you are sensing this one it could be sensed in a distorted fashion also depending upon the surface it may not be very straight. So, from this point; this point is calibrated with respect to that this corresponds to this one which has been predetermined which has been known to us from the system imaging system itself and this point corresponds to this one. And if you can interpolate this path then you can get for every point they corresponds to particular point in that light strip.

So, you need not project multiple times only from single projections you can get information of all the surface which is surface points. So, which is lying on this light stripe. So, this is called structured light range sensing system because the light itself has been structured in different patterns this is a very familiar pattern vertical stripes are very familiar pattern. And in this pattern the 3-D position of projected ray is already encoded and those encodings are used to determine the surface points. So, this is a summary, so you should get 3-D positions of all the scene points lying on the projected strip by this mechanism.

(Refer Slide Time: 13:11)

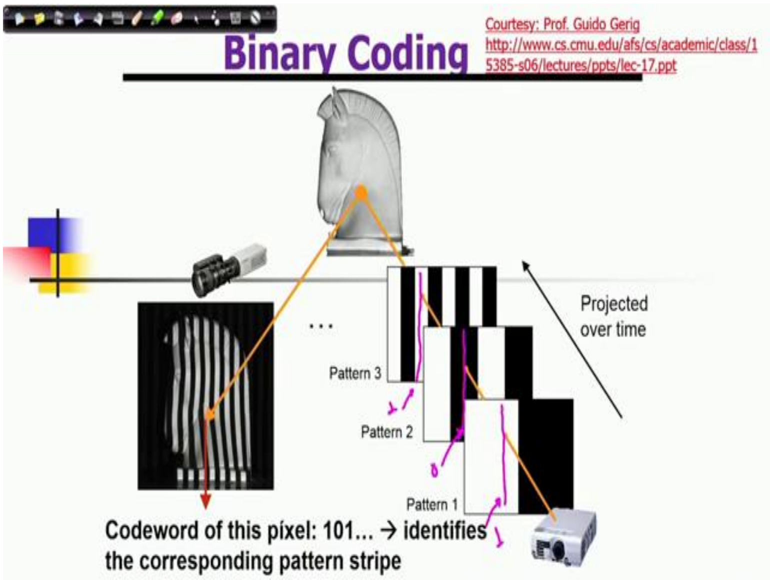


So, another way by which the structured lighting could be better utilized and where instead of projecting, but only single vertical stripes once again by scanning over the surfaces and you require more number of projections in that case. But you can encode a vertical strip position of a vertical strip in terms of m projected patterns.

So, each pattern will try to identify a zone of the surface and finally, the by observing whether a pixel is illuminated or not illuminated by the corresponding you know projection rays in m patterns by observing those strips, then you will get 2^m strips. And in this way you can get 2^m stripes and using just m patterns you can have 2^m I did not distinguish 2^m stripes on the projected surfaces. And for each one you can again apply the law of triangulations as we discussed for the structured lighting.

So, just an example. So, you have number of stripes in this pattern usually it gets doubled at every new pattern this is a particular mechanism particular type of structured light patterns and each light point is associated with an m-bit binary code. And its image is observed in your in the scene in the camera and from there you can solve the triangulation losses.

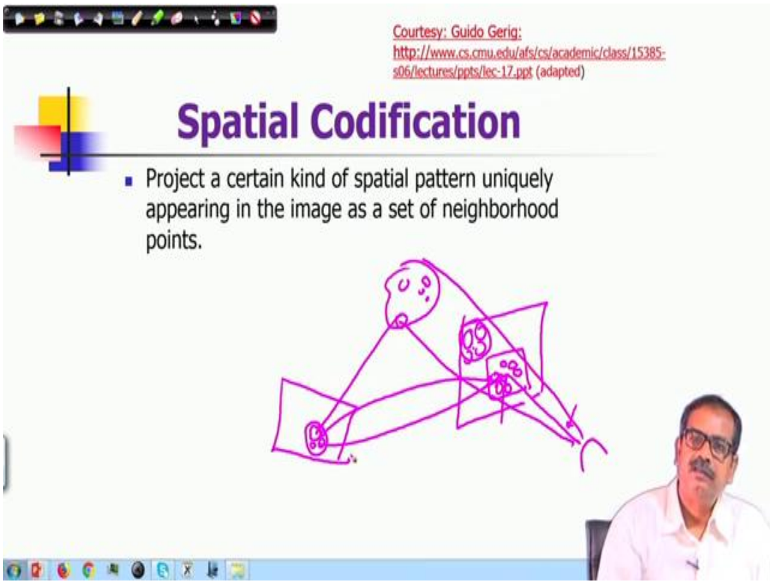
(Refer Slide Time: 14:55)



So, this is this diagram is showing how actually this panel this any particular stripe is encoded using m pattern as you can see here that : consider this ray and consider a vertical striped here. So, this strip is encoded so, in this location pattern say it is illuminated is 1 whereas, in this pattern this particular strip is 0 once again this is 1.

So, if you are using m pattern you will have m such stripes and that would give you a binary string of length m and so the you will get 2^m such discrimating patterns. So, code word could be 101 and it goes depending upon the kind of pattern you are placing there and that identifies this corresponding stripe. And from the stripe once again using the triangulation rule you can get all the depth informations for all the scene points which are lying on that particular stripes on the surfaces.

(Refer Slide Time: 16:10)

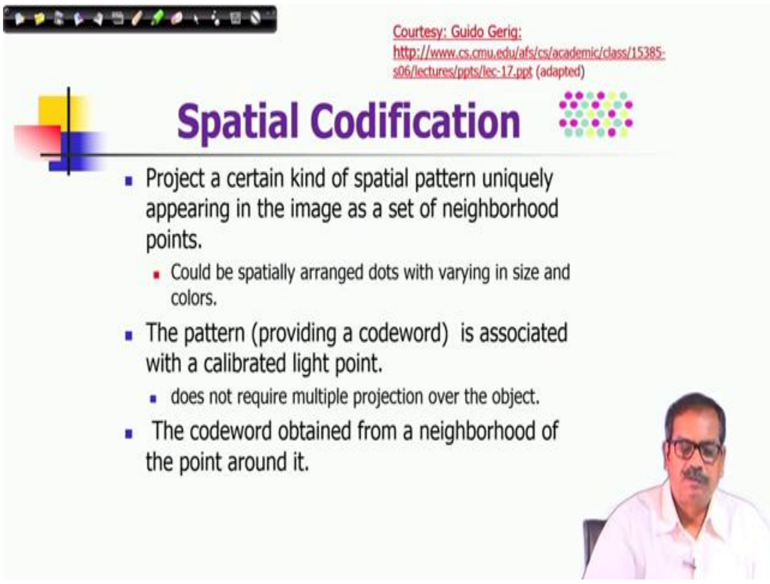


Now, there are other kind of codifications of you know spatial or the direction of the projected ray where you can consider that your and in this method you did not project it you know multiple times. So, what you can do? You can just simply illuminate only once, but there is a variation of spatial pattern on the greed of projected rays. So, every variation every and it is uniquely appearing at a particular location. So, which means if I consider an image suppose you have a projected ray and as I was mentioning that you are trying to identify the direction of the ray by a calibrated light source.

So, this location is modulated by different shapes of patterns through which this lights are projected. So, those will be visible on the surface of the visible means those can be detected it may not be visible in naked eye, but those could be detected by the camera sensors this patterns could be detected which varies in shape and also in colour and also in size. And this is uniquely identifying this particular pattern for example, it uniquely identifies the central point. So, in other locations you will have other kind of variation. So, each pattern in a location will be uniquely identified this location.

So, there is a resolution spatial resolutions by which you can encode this directions that is the limitation of this technique. But; however, the advantage is that only single projection can give you all the encoded ray which is illuminating the object. So, in the object you are having all the encoded ray simultaneously and those patterns are projected and from observing in your image if I observe those patterns wherever it is. So, since the pattern. So, you just find out the matching of this pattern with respect to your expected pattern and the this pattern library then you know that this is a image of this location. So, and then you can apply the triangulation.

(Refer Slide Time: 18:53)



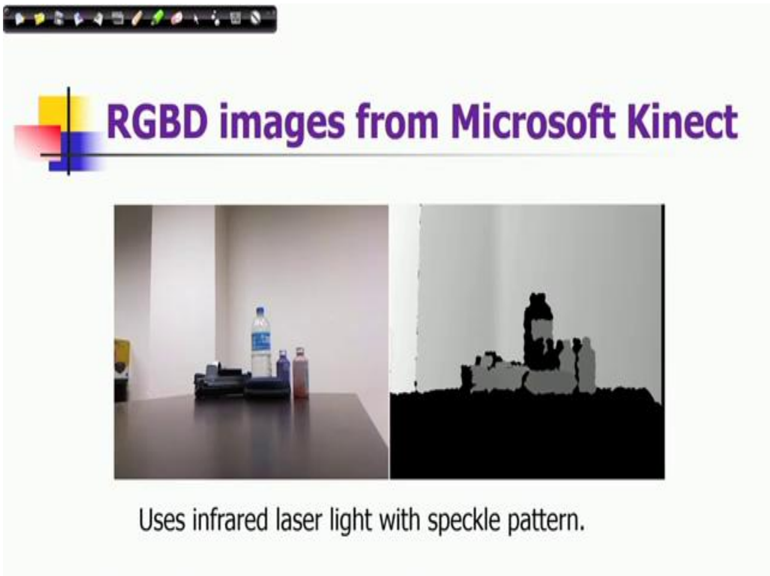
So, this technique is very effective it makes the imaging very fast and also your imaging system also becomes cheaper following this technology the resolution of the images will be less as you can understand. So, as I mentioned, as a discuss this it could be spatially arranged dots with varying in size and colors.

So, this is one example of the spatial patterns which uniquely characterizing the neighborhood. For example, if I consider this location these neighbouring locations maybe unitary you would not get any other locations which you have similar arrangement of colors or colored dots.

So, there are tricky ways by which you can design this patterns and when you detect the corresponding patterns in the camera you are also detecting that particular pattern and that it is a image of this location. So, this location is already encoded there is a particular location or direction of the ray which is given by once again two indices of their plane of projection if I say where the rays are discretized and using then you can apply; you can apply triangulation to solve it.

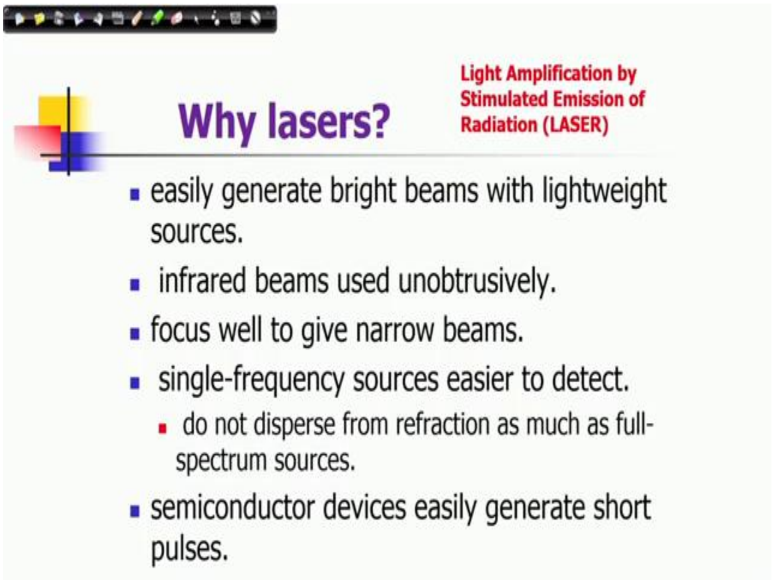
So, it provides a code word is associated with the calibrated light point and it does not require multiple projection over the object that is the advantage. And this code word is obtained from a neighborhood of the point around it that I explained also while discussing its principle.

(Refer Slide Time: 20:37)



So, one example once again the same example of range image I am showing here because this particular sensor which is Microsoft Kinect sensor, it uses infrared laser light with speckle pattern to get the same surface points and to encode the corresponding light points as I mentioned using spatial patterns. You can see that both the images are captured by the same camera, it has the corresponding optical image. So, Kinect also has an optical camera and along with there is a range camera range depth sensors following this which is captured following this principle. So, that is why the images what you are captured by Kinect is known as RGBD images, here RGB stands for red green blue components of the optical image and D stands for depth for the range image.

(Refer Slide Time: 21:39)



So, when we discussed about this techniques, we mentioned that mostly the light source what is used for this kind of projection is lasers. Now, you know the full form of LASER that is Light Amplification by Stimulated Emission of Radiation. Now, why lasers are used? You could have used any ordinary light source also for the structured light and for the triangulation schemes particularly of course, for time of light you require a coherent source where lasers have to be used.

But in other cases also lasers are mostly used compared to with respect to the ordinary lights. The reasons are that first that it easily generate bright beams with lightweight sources. So, that is one advantage of lasers. Then it can generate also infrared beams and infrared lasers are used for generating those beams and which could be used unobtrusively which means that a viewer would not be disturbed by those speckle patterns. Viewer will not get distracted or normal activities will not be hampered or will not provide the disturbing informations or abstract the viewing of the objects are seen.

So, that is why this the infrared beams are mostly used for generating this patterns nowadays particularly in Microsoft Kinect it is used and in actions like gaming etc Kinect camera could be used very easily without disturbing the participant or viewer.

For lasers also this is another good feature that it focus well to give narrow beams. And also the frequency of the laser that is also single frequency and it is easier to detect that is one advantage. It does not disperse from refraction as much as full spectrum sources; so this is another advantage. And there are semiconductor devices which easily generate the short pulses.

So, this is a reason why lasers are used instead of using any other ordinary light source in the rain sensors particularly the sensors where there are options that could have used other light sources still lasers are used in structured light based sensors that we discussed.

So, let me give a break here and we have so far we have discussed about different sensing mechanism, and from the next lecture we will be discussing about different theory and fundamentals for processing range data. Particularly will have a brief discussions on different concepts of differential geometry which are useful for characterizing the surface points of the range images. With this let me stop here.

Thank you very much for your attention.