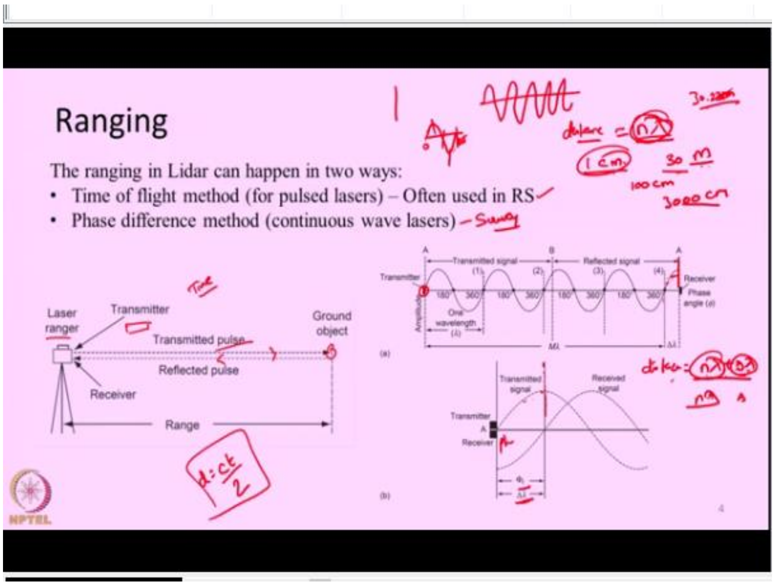
Remote Sensing: Principles and Applications
Prof. R. Eswar
Department of Civil Engineering and Interdisciplinary
Program in Climate Studies
Indian Institute of Technology-Bombay

Lecture-57
LIDAR-Part-2

Hello, everyone welcome to the next lecture on the topic of LIDAR remote sensing. In the last lecture we got introduced to what LIDAR remote sensing is and its basic working principle.

(Refer Slide Time: 00:31)



I basically told you LIDAR is a ranging tool which measures distance between the transmitter and the target of interest. So, essentially the heart of LIDAR remote sensing is ranging, how precisely or how accurately we range the distance between the transmitter and the receiver. This ranging can happen in 2 ways, one is the time of flight method and the other one is the phase difference method.

So, time of flight method which is often used in remote sensing is the simple way. We know light as a laser beam will be transmitted towards the target, it will be reflected off the target and it will come back. So, this particular transmitter will measure the time taken for the beam to go and come back. And using like simple principle of d = ct/2, where c is the velocity of light or electromagnetic radiation, t is the time taken for this entire two way flight and divided by 2 we do it because the beam has to go in the forward direction and again come back towards the receiver.

So, using the simple formula we can calculate it. This is the time of flight method and most often used in remote sensing. In some survey instruments we also use a phase difference method. Phase difference method means we send in continuous beam of laser. So, here we send a pulse, one pulse will be transmitted for certain duration say 1 nanosecond, 5 nanoseconds like that.

For a very short time period a pulse will be transmitted towards the target which will go reflect and come back, after that we will send in another pulse. But in case of continuous beam it will be continuously transmitting laser wavelength with the carrier wave, we will transmit it, it will go and come back. So, based on the distance between the transmitter and receiver the phase of the wave will vary. Let us say at this point we are transmitting a wave we know the phase at which it is transmitted. And based on the distance, the phase will change say. If the distance is integral multiple of wavelength then phase may not change. For example let us say the wavelength is 1 centimeter, say the distance between the transmitter and receiver is 30 meters. So, 30 meter is kind of an integral multiple of this, say every meter has 100 centimeters.

So, this is 3000 centimeters which is an integral multiple, if the distance is exactly like this, the phase will not change because whenever it travels a distance of 1 wavelength it will complete 1 full cycle of phase 0 to 2π. 0π/2π, 3π/2 and 2π. So, one full phase cycle is completed. So, if the distance is integral number of wavelength the phase will not vary.

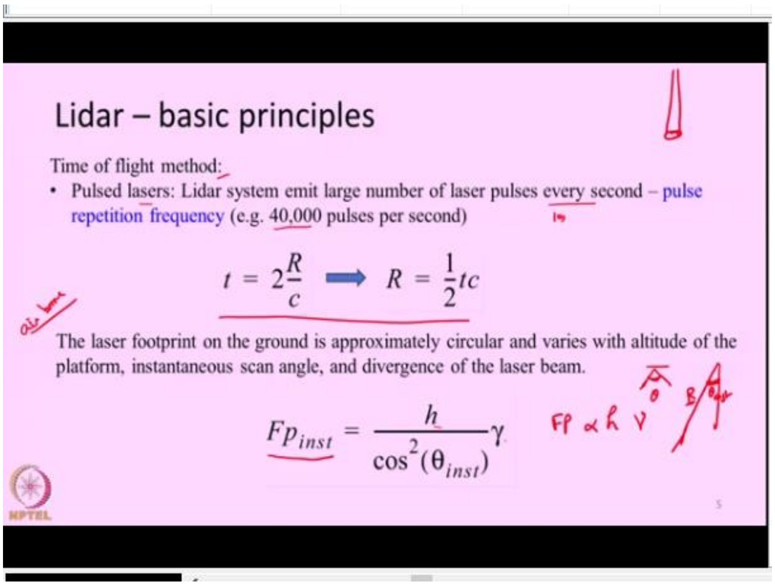
But if the distance is even slightly different than this, integral multiple of this particular wavelength then the phase of the received signal and transmitted signal will vary and by applying this phase relationship or by observing this phase relationship, we can calculate the Δλ or the small difference in distance between them. And by independent methods we can measure this nλ. So, the total distance is equal to nλ + Δλ, where nλ is measured separately. The integral multiple of wavelength plus the Δλ. Let us say the distance is 30.2 meters. If the distance was 30 meters in our example of 1 centimeter it will be nλ only will be there, this Δλ will not be there. If the distance is 30.22 meters or 2 to 5 meters which is not integral multiple of 1 centimeter, this Δλ has to be measured separately. So, that measurement will occur separately and they will be added together. This is using phase difference, but this is normally not being done for airborne or space-borne LIDAR systems, there we use the time of flight method.

We observe the time taken for the laser pulse to go and come back and use that. So, essentially in the time of flight method the entire accuracy of the ranging system depends on the accuracy with which we measure the time taken, that is one thing. We have to precisely measure the time in the order of like picoseconds. So, we can measure even very tiny fraction of time, that is possible, because light travels with very huge velocity 3×10^8 m/s. So, even in 1 nanosecond difference we will be making errors in the distance measured. So, that time should be measured very precisely and accurately. And also we need to know how a clear knowledge about the velocity of the laser beam.

If it is a short terrestrial application, say I am here, the target is in front of me, the distance is short, so not much of atmospheric change will happen, velocity I can consider as constant, then it is fine. But let us say I am doing it from space, so when the laser beam travels through space and there can be different, different mediums as we know the atmosphere is kind of stratified, it has different layers. So, each layer has certain property, the velocity of light will change when it passes through different medium, it has to pass through ionosphere, troposphere, mesosphere, lot of different layers, the velocity may change. So, we need to know accurately how the velocity change and people will measure them or model them in order to get an accurate distance measurement.

So, the major factors we need to keep in mind in the time of flight method is the measurement of time, it should be precise and accurate and also we should have the knowledge about the velocity of light as it travels through different medium.

(Refer Slide Time: 07:20)



So, this slide contains some basic formula for the time of flight method. So, a pulsed laser is what they use? They send in laser pulses, it emits large number of pulses every second, maybe some ground-based systems will be emitting 10000, 40000 pulses per second, some satellite based system will be in the order of say 10000 pulses per second and so on. And then the time is measured very simply using the range formula that we have already seen. Let us say we have an airborne platform with a scanning mechanism attached to it, it can scan certain angle of θ, at that circumstances the footprint size of the laser beam can be calculated as

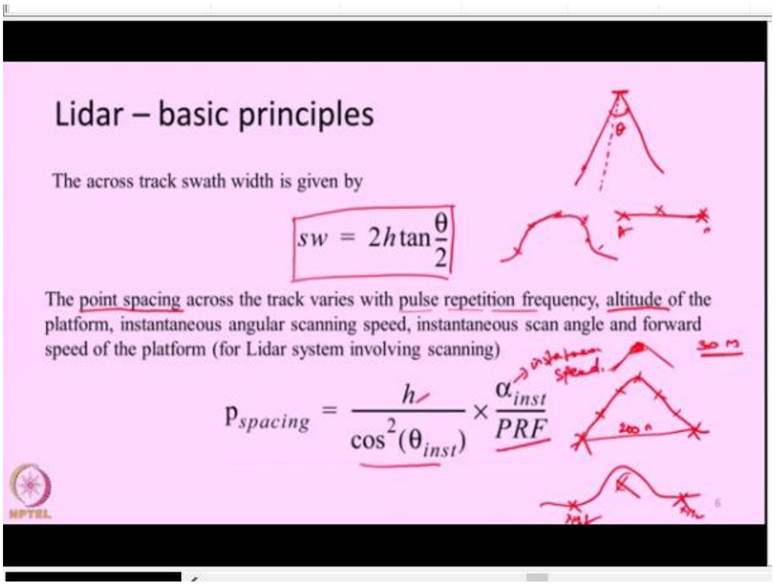
$$F_{P_{inst}} = \frac{h}{cos^2(\theta_{inst})}\gamma$$

So, footprint is kind of like the projection of laser beam onto the ground. Say whenever a laser beam starts, it will directly come in and heat the surface. So, it is kind of projection of the beam that originated from the transmitter with a minimal divergence. So, it would not be coming in perfectly parallel, there will be some small divergence. So, based on the divergence and the flying height the footprint size will vary essentially.

So, the footprint basically depends on the flying height and the divergence of the beam. This is applicable if the beam is transmitted towards nadir, but if the aircraft has a scanning mechanism attached to it then as the scanning mechanism moves away from nadir then the footprint is more or less circular in nadir which will project a small circular area on the ground. So, based on the angle, it may become slightly elliptical and the footprint size will increase. In order to account for that change we are using this $\cos^2\theta_{inst}$, where θ_{inst} is the instantaneous scan angle. Let us say this is the position at which the measurement was made. So, the scan angle at this particular instant at which this range r is measured is θ instant.

So, there can be many different θ instances for each location, for each scan position. So, this is with respect to airborne platform. Normally the space-borne platforms that are currently in operation they do not do any kind of scanning, they just send in pulses at nadir. So, this would not happen. This $\cos^2\theta_{inst}$ may not be taken care of and the laser pulse depends on the flying height and the divergence of the beam. Basically it produces a circular footprint on the ground. And similarly if your system is attached with a scanning mechanism you can define a swath width; again it is very similar to what we have learnt already for our normal whiskbroom scanners, the swath width is given by 2htan(θ/2) where θ is the entire angle with which the system can scan.

(Refer Slide Time: 10:34)



Whereas the entire θ is the total scan angle at which the systems can scan. They will certainly have multiple beams that are oriented in different, different distances and it will measure the range of those footprints that are separated by certain distance. And as the satellite moves in different, different orbits it will cover the entire ground. Then one more important thing is the point spacing in both along track and across track direction. What is this point spacing? Laser beam is a ranging device we know. So, for each point at which it hits the ground we will measure the range and effectively we can calculate the coordinate x, y z of that particular point. So, if we want to measure the terrain in a very precise way we need to have a large number of points. Let us say we have a small hill like this. So, the flight is flying into the direction of the board or your computer screen. So, this is the across track direction.

So, it is doing some sort of scanning. Let us say the point spacing happens once every 300 meters very coarse but just for explanation let us say this hill has a base distance of say 200 meters. So, one point may be collected here, another point may be collected here or maybe somewhere at a distance. So, you will know the x, y, z of this particular point, but we are going to miss this small mound in between. When we process that we will join them using some sort of interpolation mechanism. So, then our interpolation mechanism is not going to reproduce the mound or the small hill in between it is going to interpolate this point A and point B, the mound is missed.

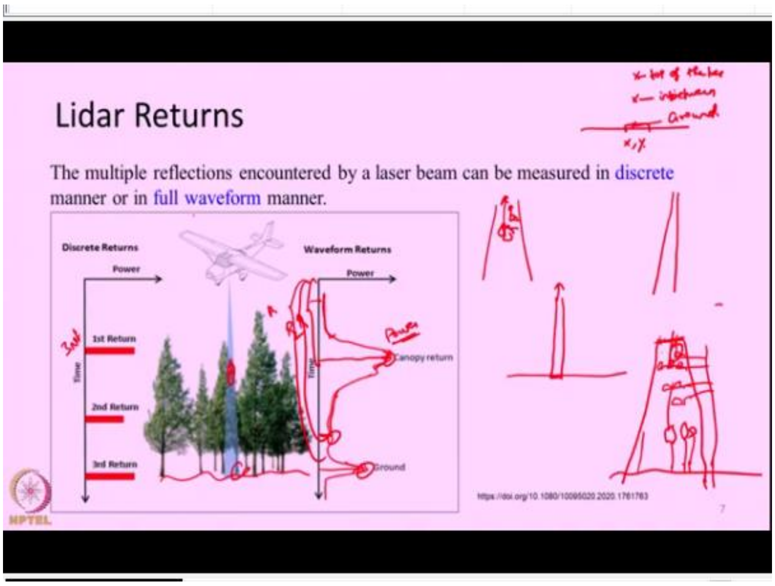
So, essentially for us to get a proper representation of the terrain it is normally to look for a high point density. Let us say the point density is pretty high, we are measuring several points along the way or in the across track direction. If that is the case then we will be able to measure the x, y, z of all the points and our interpolation mechanism will properly interpolate the surface, we will be getting a more accurate representation of the terrain.

So, the point spacing basically defines how dense or how sparse we collect the ground base observations. So, this is also one of the important parameters we should keep in mind when doing LIDAR surveying. That is why terrestrial systems will have a very high point densities, it will try to cover almost lakhs and lakhs of points. If you want to cover a big building it will send like enormous amount of pulses to get the signals back. So, the point density will determine how precisely we are able to track the surface and map the minor undulations present within it. So, this point spacing in the across track direction varies with pulse repetition frequency. Pulse repetition frequency is how frequently our system will transmit LIDAR signal, may be 10000 pulses per second and so on. And then altitude of the platform at which height the aircraft is flying, then the instantaneous angular scanning speed that is at which speed we are scanning, more the scanning speed we are going to have a lower point density, then instantaneous scan angle, forward speed of the platform. All these things are going to affect our point density.

So, for an airborne system the point spacing is given by the altitude of the platform, $\cos^2\theta_{inst}$, α_{inst}. α_{inst} is the instantaneous angular scanning speed and PRF is the point repetition frequency or pulse repetition frequency. So, this is one of the important parameter like wherever points are spaced that is important to know. Till now we are discussing about the LIDAR transmission, how LIDAR will receive the returns and how the signals may be recorded. Now we are going to see, within a footprint, there need not be only one feature with uniform elevation. There can be several features.

Let us say you have bare ground without any undulation. Then the single footprint of laser beam is going to hit this particular surface and it is going to go back. On the other hand let us say you are going to measure something over a land surface covered with trees or vegetation and there can be like small, small leaves present within it. Let this be the beam of the laser, this is the terrain. So, within this beam of laser there will be light photons everywhere present continuously.

(Refer Slide Time: 16:38)



So, within the beam there are many different scattering elements present, say this leaf will reflect some portion back, this leaf will reflect some portion back and this plant may reflect something back, there can be some reflection happening from the ground. So, essentially the returns will be not a single return but it will be kind of a continuous wave. So, with respect to the time it will vary. So, this maybe the top single leaf, it can be a large canopy, then these can be small, small leaves, then again this is the ground. So, there will be multiple returns per pulse of laser light, when there are more number of scattering elements present within the beam arranged at different, different elevations.

If that happens the laser return will be in terms of a continuous signal, like there are many different returns going back, not all laser systems will be able to record all the returns that is coming back. This incoming returning signal can be recorded in 2 ways, one is called a discrete return and another is called the full waveform return. A discrete return means the systems can store a certain number of return points per pulse, say 2 returns per pulse, 4 returns per pulse and so on. Whereas some systems can store the entire signal that came back, say if there are 50 or 60 reflections happened within that small laser beam, all the 50 or 60 can be stored by some system. So, the one with discrete return will store only a selected portion of signal. Whereas a full waveform LIDAR system can store almost the entire return that came back.

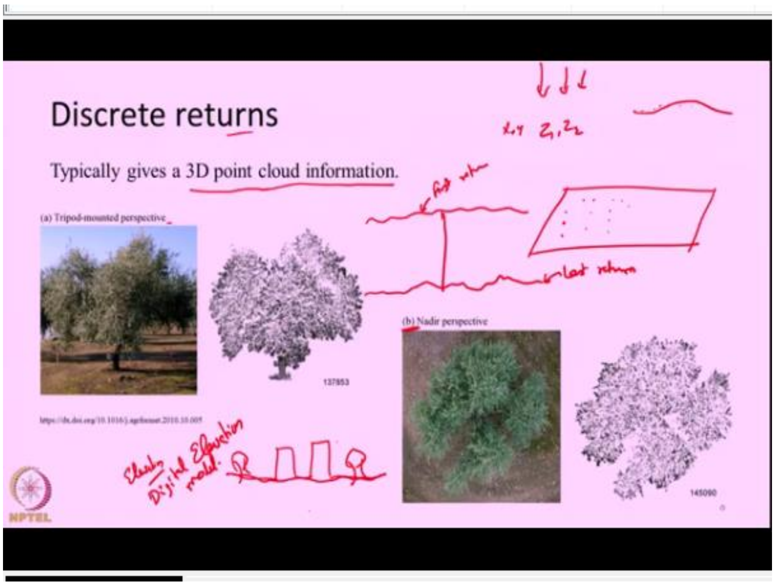
So, how the discrete system will work? Say this is the power of the signal that came back. So, first thing is with respect to time we will be able to know what elevation it came in. Based on the scattering happening each portion will return a small fraction of the incoming power, like a single leaf may produce a very weak return. So, naturally bunch of leaves will produce a higher return or more power will be reflected back when you compare this with single leaf. So, with time, with difference in arrangement of different scattering elements in the elevation, in the z direction, the power also will vary. So, when there is large number of scattering elements with high backscattering capacity, high power will be returned back.

Let us say our system is capable of storing 3 returns. So, this is one instant where there was high power return. So, store the power and the time there. This range will be measured, this power will be measured. Then the next power came in here somewhere. So, this range will be measured as R_2, this power will be measured. Then the third large signal came from the ground. So, this range again will be measured as R_3 and this power will be measured. So, basically it takes the instances where multiple power came in within that entire return signal. It will sample only those points where it got multiple power returns.

So, it may store 2 points, 3 points or 4 points based on the system. So, for each laser pulse we will have multiple ranges measured along with its power or intensity. So, for that particular ground point say x,y, you will have 3 different points for that 1, 2, 3, this can be from the top of the tree, this can be in between maybe like small leaves, this is from the ground. So, we have 3 different elevation measurements for that particular footprint. For that x, y if you consider one ground point you have 3 different measurements. Normally when you use this for a large two dimensional area we will have a point cloud. So, this is the discrete way of measurement of LIDAR signal.

This was the earliest developments and most of the systems worked in this discrete mode, but people realized for some of the applications especially for vegetation monitoring, storing the full return is beneficial for several modeling applications and so on. So, some of the LIDAR systems now stores the entire signal that is coming back either in analog form or even in digital form with a very high sampling frequency. So, it will try to store as much as high number of observations coming in within the beam. Say some systems may store 50 or 60 returns sometimes which will be enough to cover almost all normal layers 50 or 60 returns within each pulse that is possible. So, a full waveform LIDAR may produce more or less continuous representation of the signals that is coming back. The entire waveform that got reflected can be stored in the LIDAR system. So, LIDAR system can be discrete or full waveform.

(Refer Slide Time: 24:44)



So, a discrete return system will produce a 3D point cloud information, that is for each x, y you will have many number of z points. If you are talking in terms of elevation measurement or in terms of terrestrial perspective where you are measuring horizontal distances, each point is kind of a scattering element which produced a back scatter towards the target. So, from nadir perspective, if you look in three dimensions say for each x, y you will have many number of z, it will appear kind of a point cloud.

First return means whatever came in first from each pulse you can save it together, then you can store second return then you can store third return. So, if your system is capable of storing 3 returns per pulse you can store them separately as 3 different layers first return, second return and third return. And when you see them simultaneously you will get some kind of picture about the three dimensional nature of the terrain.

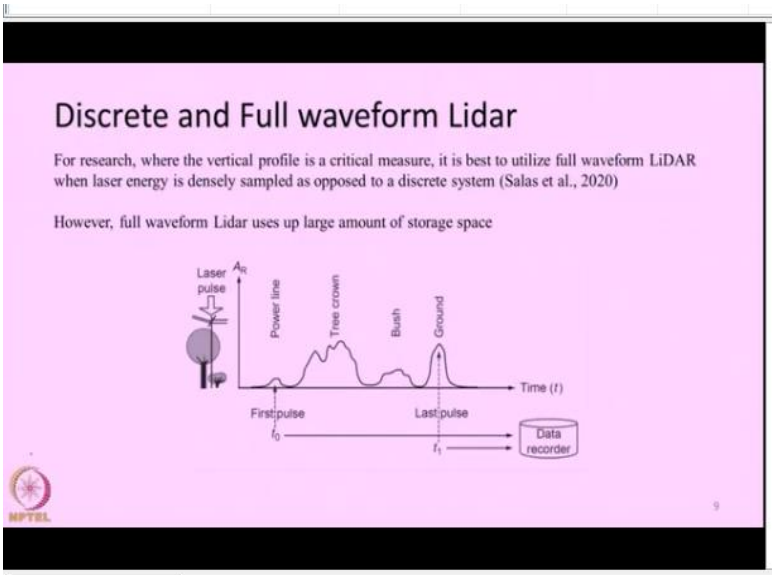
Say if you are running the flight over a forest, all the first return might have come from the canopy. Then all the last returns could have come from ground. So, if you have n number of points in one particular line you can join them using some interpolation mechanism. So, this is the interpolated first return point, this is the interpolated last return point. So, at any given point the difference in first return and the last return may give you the height of the structure that is standing over the point.

Say if it is a tree, if the first return came from the top of the tree and if the last return came from the ground then last return minus first return is going to give you the height of the tree. Imagine if you are able to do this over entire 2D space we are going to get a kind of a three dimensional representation of the topography. So, this is the major advantage of using a LIDAR system. Instantly you will be able to generate a point cloud, if you are able to filter the point cloud properly apply some processing algorithms to it, there are highly specialized algorithms to it, we are not going to see them in detail but just I am telling you if you are able to process this and separate this layers, we can do some sort of interpolation to it and the topography of the terrain with all the features attached with it.

Let us say I need to have a flood model or I need to model how if at all there is a rain how the flood water will recede? for which surface topography is one of the very important information. If that is the case then let us say I have a small a urban settlement with buildings, trees and so on. So, if I have all the first returns and last returns then I can create a elevation layer called digital elevation model which will tell me what is the elevation of my bare earth with respect to some datum or what is the elevation of my entire topography including the trees, buildings and so on. All these information will be able to get from by processing this point cloud.

So, all the first returns might have come from the top surfaces, buildings or tree tops, all the last returns might have been from the background surfaces and so on. We will be able to model the elevation of the terrain; how the topography looks and all these things using this discrete return LIDAR system.

(Refer Slide Time: 29:24)



But as I told you for some research purposes especially lie for vegetation monitoring, carbon cycle applications, vegetation biomass application it is always recommended to store the full return of the waveform that came back. So, full waveform LIDAR is preferred whenever you want to measure biomass of a tree standing especially forested trees they may have large number of leaves which will produce multiple returns. So, if you store all of them we will be able to position or we will be able to come to some sort of idea on how many returns came in. So, that many scattering elements were present within the tree, it can be many number of leaves. So, from this indirectly we can model the biomass, all these things is possible with full waveform LIDAR which is not possible from discrete return LIDAR.

So, full waveform LIDAR also has received lot of attention nowadays and people are using it for several applications but one thing we have to remember is a full waveform LIDAR will use a lot of space in the computer memory. Just imagine instead of taking 4 samples per pulse now we are taking 60 to 70 samples per pulse which is a tremendous increase, so that will have a large strain on the system. So, normally a full waveform LIDAR will use a lot of storage space.

So, as a summary in this lecture we discussed about further basic principles of how LIDAR system works and also we discussed about a full waveform LIDAR and a discrete return LIDAR. With this we end this lecture.

Thank you very much.