Introduction to LiDAR Remote Sensing

Cheng Wang, Xuebo Yang, Xiaohuan Xi, Sheng Nie, and Pinliang Dong

First published 2024

ISBN: 978-1-032-67150-5 (hbk)
ISBN: 978-1-032-67152-9 (pbk)
ISBN: 978-1-032-67151-2 (ebk)

Chapter 2

LiDAR Remote Sensing Principles

DOI: 10.1201/9781032671512-2

高等教育出版社
HIGHER EDUCATION PRESS

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

2 LiDAR Remote Sensing Principles

2.1 LiDAR RANGING PRINCIPLE

Ranging is a very important application of LiDAR systems. There are two main LiDAR ranging modes: pulse-based ranging and phase-based ranging. The ranging principles of the two modes are described in the following sections.

2.1.1 Ranging Principle of Pulsed LiDAR

At present, most LiDAR systems work in the pulsed-based ranging mode. That is, the laser emits a very narrow pulse (the pulse width is usually less than 50 ns) toward the target, and then part of the pulse is reflected by the target and returns to the receiver. The LiDAR system records the time of flight (TOF) between the emission and reception of a laser pulse and then calculates the distance between the senor and the target according to the speed of light, as in Equation (2.1). Figure 2.1 shows the ranging principle of pulsed LiDAR.

$$R = \frac{1}{2} \cdot c \cdot t \tag{2.1}$$

where R is the distance between the sensor and target, c is the speed of light in air, and t is the TOF of the laser between emission and reception.

Differentiating Equation (2.1), we have Equation (2.2):

$$\Delta R = \frac{1}{2} \cdot c \cdot \Delta t \tag{2.2}$$

where ΔR is the ranging resolution, representing the minimum distance at which two objects can be distinguished, depending on the time sampling interval.

Additionally, in order to ensure that the returns of multiple laser beams can be distinguished, the pulsed LiDAR system must transmit the next laser pulse after receiving the return of the previous laser pulse. In this case, the maximum measurement distance (R_{max}) needs to be considered (Lai, 2010). It is calculated by the longest time measurement (t_{max}), as in Equation (2.3):

$$R_{max} = \frac{1}{2} \cdot c \cdot t_{max} \tag{2.3}$$

In practical applications, the maximum measurement distance is affected by many factors, such as laser power, beam divergence, atmospheric transmission,

DOI: 10.1201/9781032671512-2

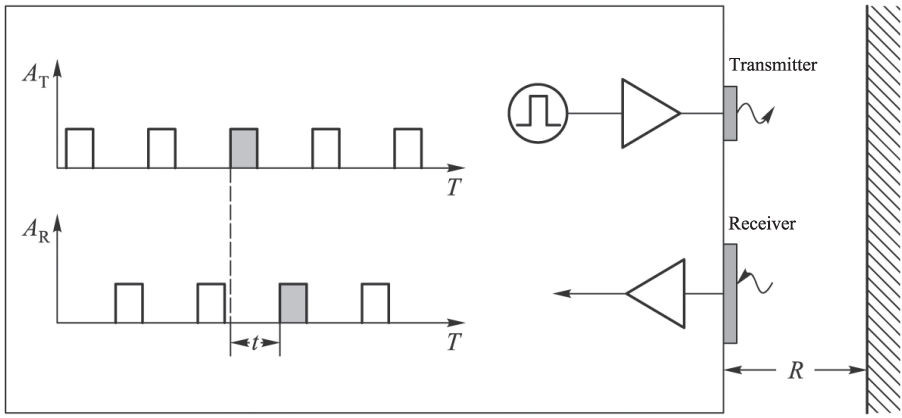

FIGURE 2.1 Ranging principle of pulsed LiDAR.

target reflection characteristics, detector sensitivity, flight altitude, and deviation of flight attitude record.

2.1.2 Ranging Principle of Phased LiDAR

The phase-based LiDAR is also called continuous wave LiDAR. It transmits an intensity-modulated continuous laser which is reflected by the target and then returns to the receiver. The distance between the target and sensor is calculated by the phase variation of the laser during the round trip, as in Figure 2.2. Generally, the ranging precision of a phased LiDAR system is higher than that of a pulsed LiDAR.

Assuming that the modulated continuous laser wave is a sine waveform with a period of T, the ratio of the time interval of the measured phase difference to the

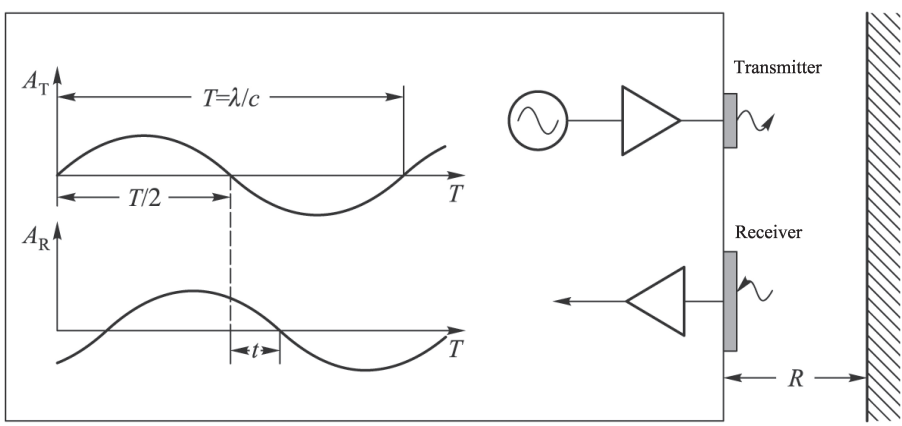

FIGURE 2.2 Ranging principle of phased LiDAR.

period T is exactly the ratio of the phase difference of the transmitted wave and the received wave to the phase 2π, as in Equation (2.4).

$$\frac{t}{T} = \frac{\phi}{2\pi} \tag{2.4}$$

where ϕ is the phase difference of the transmitted wave and the received wave. According to Equation (2.4), the distance between the sensor and target (R) is expressed as in Equation (2.5).

$$R = \frac{c}{2} \cdot \frac{T}{2\pi} \cdot \phi \tag{2.5}$$

where the period T is the reciprocal of the modulated frequency f of the sine wave. In this case, R can be expressed as in Equation (2.6).

$$R = \frac{1}{4\pi} \cdot \frac{c}{f} \cdot \phi \tag{2.6}$$

where the ratio of the speed of light (c) to the modulated frequency (f) is exactly the wavelength (λ). Hence, the distance R is as in Equation (2.7).

$$R = \frac{\lambda}{4\pi} \cdot \phi \tag{2.7}$$

Differentiating Equation (2.7), we get the ranging resolution (ΔR) of phased LiDAR, as in Equation (2.8).

$$\Delta R = \frac{\lambda}{4\pi} \cdot \Delta\phi \tag{2.8}$$

Equation (2.8) indicates that the ranging resolution of the phased LiDAR is related to not only the phase resolution but also the laser wavelength. In contrast, the ranging resolution of the pulsed LiDAR is only related to the time resolution [Equation (2.2)]. With the same phase (time) resolution, the ranging resolution of the phased LiDAR system could be improved by using a shorter wavelength. Hence, the phased LiDAR is usually used in applications with high ranging precision. For example, for a continuous laser wave with the period of 1 second, assuming that the phase resolution is 0.4°, its ranging resolution is 0.1 mm. However, to achieve the same ranging resolution, the pulsed LiDAR system must have a time resolution of 1 picosecond, which requires higher standards for the time interval counting hardware system.

Although the phased LiDAR system is more suitable for high-precision ranging applications, it is usually used for short-range detection. This is because it requires high power to emit a continuous laser. Also, it may cause ranging ambiguity. The maximum measured distance (R_{max}) of the phase-based LiDAR system which can

uniquely determine the target is the distance measured when the continuous wave travels exactly one period, as in Equation (2.9).

$$R_{\max} = \frac{\lambda}{4\pi} \cdot \phi_{\max} = \frac{2\pi}{4\pi} \cdot \lambda = \frac{\lambda}{2} \tag{2.9}$$

Equation (2.8) indicates that the ranging resolution (ΔR) of the phased LiDAR system is related to the phase resolution ($\Delta\phi$) and the laser wavelength (λ). The shorter the laser wavelength used, the higher the ranging resolution. Hence, phased LiDAR makes it relatively easy to achieve high ranging resolution. However, Equation (2.9) shows that the maximum measured distance ($R_{\max}$) of phased LiDAR is determined by the wavelength (λ). The longer the wavelength, the longer the maximum measured distance. Long-distance measurement and high ranging resolution both cannot be achieved in a continuous wave LiDAR system. Therefore, in order to meet the requirements of practical applications, the phased LiDAR system is usually equipped with a signal frequency modulation device that can adjust multiple frequencies. Additionally, high-power phased lasers are difficult to produce. Currently, most LiDAR systems use a pulsed ranging system. The LiDAR mentioned in the rest of this book refers to the pulsed LiDAR.

2.1.3 Ranging Precision of LiDAR

Ranging precision is a pivotal parameter of a LiDAR system and is different from ranging resolution. Specifically, the ranging resolution refers to the limit of the ability to measure the distance between two objects along the same line of sight, while the ranging accuracy is the standard deviation of the target distance estimation in the presence of noise. The ranging precision is related to the signal-to-noise ratio (SNR) of the returned laser signal and the emitted laser pulse (Wehr & Lohr, 1999).

Usually, the ranging precisions of pulsed LiDAR and phased LiDAR are expressed as in Equations (2.10) and (2.11), respectively.

$$\sigma_{R_{\text{pulse}}} \propto \frac{c}{2} \cdot t_{\text{rise}} \cdot \frac{\sqrt{B_{\text{pulse}}}}{P_{R_{\text{peak}}}} \tag{2.10}$$

$$\sigma_{R_{\text{cw}}} \propto \frac{\lambda}{4\pi} \cdot \frac{\sqrt{B_{\text{cw}}}}{P_{R_{\text{av}}}} \tag{2.11}$$

where $\sigma_{R_{\text{pulse}}}$ is the ranging precision of pulsed LiDAR, $\propto$ indicates a proportional relationship, t_{rise} is the pulse rise time, B_{pulse} is the noise width of pulsed LiDAR, $P_{R_{\text{peak}}}$ is the peak power of the received signal of pulsed LiDAR, $\sigma_{R_{\text{cw}}}$ is the ranging precision of phased LiDAR, λ is the continuous laser wavelength, B_{cw} is the noise width of phased LiDAR, and $P_{R_{\text{av}}}$ is the average power of the received signal of phased LiDAR.

Here we compare the performance of the two ranging modes with the assumption of measuring the same target. In this case, the received power can be replaced by

the transmitted power. That is, the peak power $P_{R_{\text{peak}}}$ of the received laser of pulsed LiDAR is replaced by the peak power of the transmitted pulse $P_{T_{\text{peak}}}$, and the average power of the received laser of phased LiDAR $P_{R_{\text{av}}}$ is replaced by that of the transmitted laser $P_{T_{\text{av}}}$. The ratio of ranging precisions of pulsed LiDAR to phased LiDAR is expressed as in Equation (2.12).

$$\frac{\sigma_{R_{\text{pulse}}}}{\sigma_{R_{\text{cw}}}} \propto 2\pi \cdot \frac{c}{\lambda} \cdot t_{\text{rise}} \cdot \frac{P_{T_{\text{av}}}}{P_{T_{\text{peak}}}} \cdot \sqrt{\frac{B_{\text{pulse}}}{B_{\text{cw}}}} \tag{2.12}$$

Assuming that:

$$t_{\text{rise}} \propto \frac{1}{B_{\text{pulse}}} \tag{2.13}$$

Equation (2.12) can be rearranged as:

$$\frac{\sigma_{R_{\text{pulse}}}}{\sigma_{R_{\text{cw}}}} \propto 2\pi \cdot f \cdot \sqrt{\frac{t_{\text{rise}}}{B_{\text{cw}}}} \cdot \frac{P_{T_{\text{cw}}}}{P_{T_{\text{peak}}}} \tag{2.14}$$

With the assumption of the pulse rise time of pulsed LiDAR of 1 ns, the peak power of transmitted pulse of 2000 W, the frequency of phased LiDAR of 10 MHz, the noise width of phased LiDAR of 7 kHz, and the average transmit power of 1 W, the ratio of ranging precisions of pulsed LiDAR to phased LiDAR can be calculated by Equation (2.14), that is about 0.0012. This indicates that although the transmit power of the pulsed LiDAR is 2000 times that of the phased LiDAR, the ranging precision of the pulsed LiDAR can theoretically only reach 85 times that of the phased LiDAR.

2.2 LiDAR RADIATION PRINCIPLE

The LiDAR system performs non-contact detection of an object's surface by transmitting laser pulses and receiving the returned energy. Its remote sensing (RS) process includes four steps (Figure 2.3): (1) transmission of the emitted laser toward the target, that is, the pulse downward; (2) interaction of pulse and target surface, including interception, scattering, absorption, etc.; (3) transmission of the scattered laser toward the sensor, that is, pulse upward; and (4) reception of the scattered laser by the receiver. Through this process, the received LiDAR signal is a comprehensive expression of the sensor configuration, the target surface characteristics, and the surrounding environment. In order to interpret the LiDAR signal, revert the LiDAR RS process, and retrieve the target information, it is necessary to understand the radiation principle of LiDAR RS. This section presents the radiation principle of the LiDAR system, which is also the core of LiDAR's three-dimensional (3D) imaging mechanism.

2.2.1 LiDAR Equation

The physical process of the laser beam penetrating the atmosphere, interacting with scatterers, and returning to the receiver is very complex. Usually, beginners would

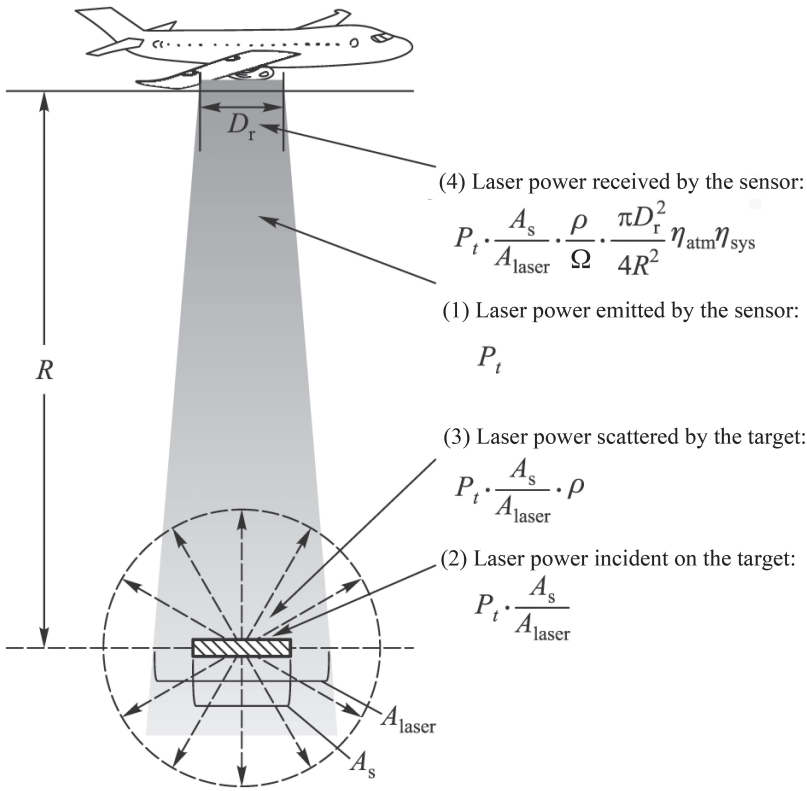

FIGURE 2.3 LiDAR RS process. (P_t: Laser power emitted by the sensor; A_s: effective area of target [scatterer]; A_{laser}: laser footprint area; ρ: reflectance of target [scatterer]; Ω: scattered solid angle of target toward the receiver; R: distance between sensor and target; D_r: diameter of LiDAR receiver; η_{atm}: effect of atmospheric transfer; η_{sys}: efficiency of sensor transmitting, receiving, and processing signals.)

simplify this complex process of LiDAR RS. The LiDAR equation is a physically based formula to compute the power returned to a receiver for given transmitted laser power, optical properties of the medium through which the LiDAR beam passes, and target properties. It is deducted based on two strong assumptions: (1) the power emitted by the laser is uniformly distributed within the footprint and (2) the scatterer uniformly disperses the incident power into a conical solid angle of Ω (Wagner et al., 2006). Based on the assumptions, the derivation process of the LiDAR equation is as follows.

The footprint area of the laser beam at the distance of the scatterer is:

$$A_{laser} = \frac{\pi R^2 \beta_t^2}{4} \tag{2.15}$$

where R is the distance between the sensor and scatterer and β_t is the tangent of the laser beam divergence angle. Based on assumption (1) earlier, the

incident laser power density of the scatterer surface S_S is expressed as in Equation (2.16):

$$S_S = \frac{4P_t}{\pi R^2 \beta_t^2} \tag{2.16}$$

where P_t is the total emitted laser power. Equation (2.16) describes the property by which the laser power density gradually decays with the transmitted distance R. Considering the reflectance ρ and the effective area A_s of the scatterer, the laser power dispersed by the scatterer P_s is expressed as in Equation (2.17).

$$P_s = \frac{4P_t}{\pi R^2 \beta_t^2} \rho A_s \tag{2.17}$$

Based on assumption (2) earlier, if the scattered conical solid angle of the scatterer overlaps with the field of view of the LiDAR receiver, the returned laser power density is expressed as in Equation (2.18).

$$S_r = \frac{4P_t}{\pi R^2 \beta_t^2} \rho A_s \frac{1}{\Omega R^2} \tag{2.18}$$

Finally, considering the receiver aperture and the influence of the atmosphere and hardware systems, the LiDAR equation of a single scatterer can be expressed by Equation (2.19) as:

$$P_{R,i}(t) = \frac{4P_t}{\pi R^2 \beta_t^2} \rho A_s \frac{1}{\Omega R^2} \cdot \frac{\pi D_r^2}{4} \eta_{atm} \eta_{sys} \tag{2.19}$$

where $P_{R,i}(t)$ represents the returned laser power from the i-th scatterer at the time t, D_r is the diameter of the LiDAR receiver, Ω is the scattered solid angle of the scatterer, η_{atm} is the effect of atmosphere transfer on laser power, and η_{sys} is the effect of the sensor transmitting, receiving, and processing signals.

All the parameters related to the scatterer can be combined into one, which is called the laser radar cross section σ (LRCS) (Dai, 2002):

$$\sigma = \frac{4\pi}{\Omega} \rho A_s \tag{2.20}$$

Equation (2.20) shows that LRCS is determined by the effective area A_s, reflectance ρ, and scattered solid angle Ω of the scatterer.

In the case where multiple scatterers are distributed in the propagated path of the laser, the received signal can be expressed as a general expression of the LiDAR equation, as in Equation (2.21).

$$P_{sum}(t) = \sum_{i=1}^{N} \eta_{atm,i} \eta_{sys,i} \frac{D_r^2}{4\pi R_i^4 \beta_t^2} P_t * \sigma_i(R_i) \tag{2.21}$$

where $P_{\text{sum}}(t)$ is the accumulation of the return signals of N scatterers at time t and the relationship of the time t and the distance R_i is $t = \frac{2R_i}{c}$, * represents the convolution operation.

2.2.2 LiDAR Waveform Model

Once there are multiple scatterers at different distances from the sensor in the laser propagation direction, the emitted laser might be reflected multiple times, and the reflected laser intensity and the traveled distance are both different for each time. The LiDAR waveform records the laser intensity according to the returned time delay, that is, laser intensity as a function of time. As in Figure 2.4, the LiDAR waveform can be quantified as a one-dimensional waveform with time as the horizontal axis and intensity as the vertical axis. It is theoretically feasible to inverse the characteristics of scatterers by accurately interpreting the LiDAR waveform. However, many parameters in the LiDAR equation are difficult to obtain accurately, such as η_{atm} and η_{sys}. Also, the parameters such as Ω and ρ are different for different objects and under different observation conditions. This makes it very difficult to directly inverse the surface characteristics from the LiDAR waveform.

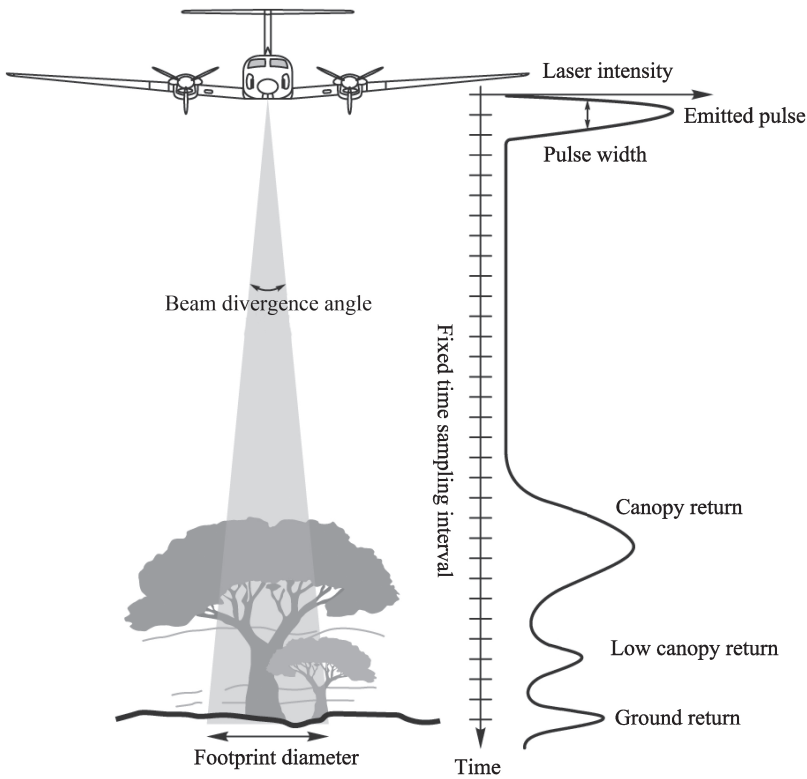

FIGURE 2.4 Schematic diagram of LiDAR waveform.

Therefore, in practical applications, the LiDAR equation need to be further simplified based on some specific assumptions.

The LiDAR waveform is a product of the interaction between the laser pulse and Earth's surface. The laser-surface interaction process can be expressed as the transfer function. Since the emitted laser pulse is close to a one-dimensional Gaussian distribution in the time domain, the LiDAR waveform is just the convolution of the emitted pulse and the transfer function in the time domain, which is usually approximated by a mixed Gaussian model. That is, the LiDAR waveform is regarded as a superposition of multiple Gaussian functions, as in Equations (2.22) and (2.23).

$$P_{\text{sum}}(t) = \sum_{i=1}^{N} f_i(t) \tag{2.22}$$

$$f_i(t) = A_i \mathrm{e}^{-\frac{(t-\mu_i)^2}{2\sigma_i^2}} \tag{2.23}$$

where $f_i(t)$ is the i-th Gaussian component and A_i, μ_i, and σ_i are the amplitude, mean, and standard deviation of the i-th Gaussian component, respectively. Based on this principle, the mixed Gaussian function can be used to decompose the waveform, and the characteristic parameters of the ground object (such as reflectivity and geometric shape) can be obtained by inversion. For more details on waveform decomposition, see Section 4.2 of this book.

2.2.3 LiDAR Radiative Transfer Model

In addition to the LiDAR equation and the LiDAR waveform model based on very strong assumptions, some researchers have proposed relatively complex LiDAR radiative transfer models (RTMs). The RS radiative transfer models are considered celebrated tools for modeling, validating, and interpreting the RS signals, as they can simulate the interaction mechanisms of electromagnetic waves and Earth's surface that give rise to the signals.

To date, many researchers have developed a wide range of LiDAR RTMs with diverse complexity and efficiency. According to the degree of introduction of physical mechanism, the LiDAR RTMs can be divided into semi-empirical models, analytical models, and computer simulation models (Table 2.1). The semi-empirical models use a digital surface model (DSM) or similar methods to represent the elevation distribution of the Earth's surface and simulates the LiDAR signal as a temporal summation of Gaussian profiles that are computed by convolving the target reflectance and pulse energy distribution (Blair & Hofton, 1999). However, they only consider the single scattering of the laser on the target surface. Physically based analytical models use strong simplification of landscapes and ray propagations. For example, Sun and Ranson (2000) established a 3D LiDAR waveform simulation model, which simplifies the vegetation landscape into a series of 3D turbid voxels and only considers the single scattering of laser rays. Ni-Meister et al. (2001) created a LiDAR waveform simulation model based on 3D geometric optics radiative transfer (GORT),

TABLE 2.1
LiDAR Radiative Transfer Model

Category	Model	Proposer	Features
Semi-empirical model	DSM-based model	(Blair & Hofton, 1999)	DSM is needed as prior knowledge; only single scattering is considered
Analytical model	3D LiDAR model	(Sun & Ranson, 2000)	Canopy is simplified as turbid mediums; only single scattering is considered
	3D GORT model	(Ni-Meister et al., 2001)	Canopy is simplified as single-layer and double-layer forest; the ground is a flat or a terrain with fixed slope
Computer simulation model	RAPID	(Huang & Wynne, 2013)	Radiosity method
	DART-RC	(Gastellu-Etchegorry et al., 2016)	Forward ray tracing
	FLIGHT	(North et al., 2010)	Backward path tracing
	DIRSIG	(Goodenough & Brown, 2017)	Forward ray tracing
	LIBRAT	(Disney et al., 2009)	Backward path tracing
	RAYTRAN	(Govaerts & Verstraete, 1998)	Forward ray tracing
	FILES	(Kobayashi & Iwabuchi, 2008)	Forward ray tracing
	DART-Lux	(Yang et al., 2022)	Bidirectional path tracing, efficient and accurate

which simplifies the ground as a terrain with a fixed slope or a flat surface without undulations and assumes that the ground has the same reflectance and the vegetation is a single-layer or double-layer forest. The semi-empirical and analytical models all make strong simplifications and assumptions on Earth's scenery and laser propagation process. Although computationally efficient, these models have an increasingly unsuitable accuracy for most potential LiDAR applications, such as signal modeling, parameter inversions of multiple targets, and LiDAR sensor design.

With the development of computer graphics, some LiDAR computer simulation models based on Monte Carlo ray tracing (MCRT) (Pharr et al., 2016) were proposed, such as DART (Gastellu-Etchegorry et al., 2016; Yang et al., 2022) and FLIGHT (North et al., 2010). These MCRT-based LiDAR 3D RTMs are the most commonly used and accurate models since they are more adaptive to complex scene structures and usually do not make simplifications of ray interactions. According to the direction of ray tracing, they are usually categorized as forward models if rays are traced from the laser source, backward models if rays are traced from the receiver, and bidirectional path tracing (BDPT) models if rays are traced from the source and receiver simultaneously. Specifically, the forward ray tracing models use the laser source as the starting point for ray tracing and record the signal that finally enters the receiver. The backward tracing models trace the path from the receiver to ensure that the traced signal is received. The bidirectional tracing model traces the rays from the source and receiver synchronously. The purpose of all these ray tracing

models is to generate a series of random paths that connect the source and receiver. For LiDAR systems with the received field of view larger than the illumination area (footprint), the bidirectional tracing model usually has the highest accuracy and efficiency, followed by the forward tracing models and the backward tracing model with the lowest.

Here we introduce a BDPT-based LiDAR modeling method called DART-Lux (Yang et al., 2022). The DART-Lux LiDAR model transforms the laser transport problem into an integration over all possible paths in the 3D landscape that connect the laser source and the receiver. It relies on the light transport equation, as in Equation (2.24).

$$L(r',\Omega_o) = L_e(r',\Omega_o) + \int_{\Omega} L(r',\Omega_i)\cdot f(r',-\Omega_i,\Omega_o)\cdot\left|\cos\theta_i\right|\cdot \mathrm{d}\Omega_i \tag{2.24}$$

where the exitant radiance $L(r',\Omega_o)$ at vertex r' along direction Ω_o is the sum of the emitted radiance $L_e(r',\Omega_o)$ and the scattered radiance due to incident radiance $L(r',-\Omega_i)$ along direction Ω_i, $f(r',-\Omega_i,\Omega_o)$ is the bidirectional scattering distribution function (BSDF) of the surface at vertex r', and θ_i is the incident angle between the incident direction Ω_i and the surface normal vector. Equation (2.24) can be transformed into an area integration form instead of a solid angle integration form, as in Equation (2.25).

$$L(r'\to r) = L_e(r'\to r) + \int_{A} L(r''\to r')\cdot f(r''\to r'\to r)\cdot G(r''\to r')\cdot \mathrm{d}A(r'') \tag{2.25}$$

where r'' and r are the previous vertex and next vertex of r' in the light transport process, $\mathrm{d}A(r'')$ is the area at vertex r'', and the connection function $G(r''\leftrightarrow r')$ between vertices r'' and r' is the product of $\left|\cos\theta_i\right|$ and the Jacobian term to transfer solid angle integration over area integration.

Since laser rays might be scattered multiple times in the 3D landscape, the received laser power can be represented by incrementally expanding Equation (2.25) to an infinite sum of a multidimensional integral, which can be expressed as a Lebesgue integral, as in Equation (2.26).

$$\Phi_{\mathrm{LiDAR}}\left(\tau(\bar{\tau})\right) = \int_{\mathcal{D}} f(\bar{r})\cdot \mathrm{d}\mu(\bar{r}) \tag{2.26}$$

where $\tau(\bar{r})$ is the laser travel time along path $\bar{r}$ and $\Phi_{\mathrm{LiDAR}}\left(\tau(\bar{r})\right)$ is the returned laser power at the time $\tau(\bar{r})$. $\mathcal{D}$ is the set of all paths. $\mathcal{D} = \bigcup_{n=1}^{\infty}\mathcal{D}_n$ with $\mathcal{D}_n$ the set of paths with n edges (n is also called the path depth). $\bar{r}$ is a path of the set of paths that connect the laser source and the receiver; it passes through a series of vertices in the scene: $\bar{r}\in\mathcal{D} = \left\{\bar{r}_n \middle| \bar{r}_n = r_0, r_1, \ldots, r_n;\ r_{k=0,1,\ldots,n}\in A,\ n\in\mathbb{Z}^+\right\}$. $f(\bar{r})$ is the power contribution of path $\bar{r}$. $\mathrm{d}\mu(\bar{r})$ is the area-product measure of path $\bar{r}$, $\mathrm{d}\mu(\bar{r}_n) = \mathrm{d}A(r_n)\cdot\mathrm{d}A(r_{n-1})\cdots\mathrm{d}A(r_0)$. The solution of Equation (2.26) is the simulated LiDAR signal. The DART-Lux LiDAR model uses a Monte Carlo–based BDPT

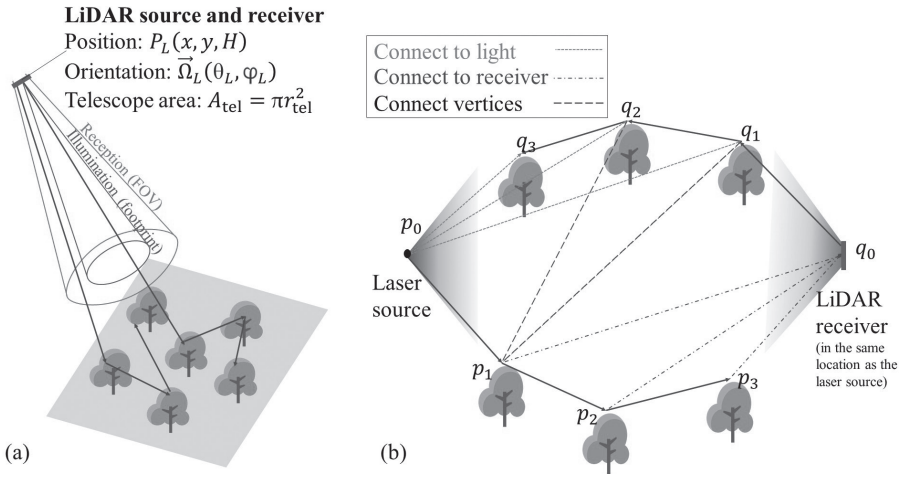

FIGURE 2.5 Schematic diagram of the DART-Lux LiDAR model. (a) A sub-path starts from the LiDAR source and a sub-path starts from the receiver (at the same location as the source). (b) A series of paths that connect the LiDAR source and receiver are generated by the BDPT method.

algorithm to solve Equation (2.26). Generally, the source and receiver of the LiDAR sensor are considered to have the same central position $P_L(x, y, H)$ and orientation $\vec{\Omega}_L(\theta_L, \varphi_L)$. The paths from laser source to receiver are constructed based on the combination of two sub-paths that start from the source and the receiver, respectively, as in Figure 2.5.

In the DART-Lux LiDAR model, the first random walk starts from the laser source and samples a light sub-path $\overline{p}_{N_s} = p_0, p_1, \ldots, p_{N_s\text{-}1}$ with N_s vertices, of which vertex p_0 is on the laser source. Each vertex is sampled with the probability distribution function (PDF) $p\left(p_k^{\rightarrow}\right)\left(k = 0,1,\ldots,N_s\text{-}1\right)$. The second random walk starts from the receiver and samples a receiver sub-path $\overline{q}_{N_t} = q_0, q_1, \ldots, q_{N_t\text{-}1}$ with N_t vertices, of which vertex q_0 is on the receiver. Each vertex is sampled with the PDF $p\left(q_k^{\leftarrow}\right)\left(k = 0,1,\ldots,N_t\text{-}1\right)$. The combination of the two sub-paths leads to the construction of a series of complete paths from source to receiver by using three sampling techniques: (1) "connect to light": the vertex on the receiver sub-path is connected to a sampled vertex on the laser source; (2) "connect to receiver": the vertex on the light sub-path is connected to a sampled vertex on the receiver; and (3) "connect vertices": a vertex on the light sub-path is connected to a vertex on the receiver sub-path. Also, a light sub-path can randomly hit the LiDAR receiver, and a receiver sub-path can randomly hit the laser source. The length $l\left(\overline{r}_{s,t}\right)$ of any path $\overline{r}_{s,t}$ is the sum of the lengths of the light sub-path, the receiver sub-path, and the distance between the connected vertices p_{s-1} and q_{t-1}. The path length is further converted to the laser travel time $\tau(\overline{r})$. The power contribution divided by the path PDF $\left(\frac{f(\overline{r}_{s,t})}{p(\overline{r}_{s,t})}\right)$ of a random path $\overline{r}_{s,t}$ can unbiasedly estimate the laser returned power Φ_{LiDAR}. The simulated temporal power profile is the so-called LiDAR waveform. Furthermore,

the DART-Lux LiDAR model extends the single-pulse waveform LiDAR modeling component to simulate multiplatform (satellite, airborne, terrestrial), multitype (waveform, discrete return, photon counting), and multipulse LiDAR signals. The multipulse simulation of laser scanning systems is an iterative loop of single-pulse modeling with specific geometry configurations and methods of detection and digitization. The point cloud of discrete return LiDAR is converted from the simulated waveform using Gaussian decomposition or peak detection methods. The points of photon counting LiDAR are derived per waveform using a statistical method based on the instrumental parameters of the single-photon detector.

2.3 PRINCIPLES OF LiDAR ON DIFFERENT PLATFORMS

According to different application requirements, LiDAR systems can be equipped on different platforms, including spaceborne (satellites and the International Space Station [ISS]), airborne (manned and unmanned aerial vehicles), ground (car, backpack, terrestrial tripod), etc. This section introduces the principles of LiDAR on different platforms.

2.3.1 Spaceborne LiDAR

Until 2022, the existing spaceborne LiDAR systems for Earth observation include Ice, Cloud, and Land Elevation Satellite/Geoscience Laser Altimeter System (ICESat/GLAS), Advanced Topographic Laser Altimeter System (ICESat-2/ATLAS), Global Ecosystem Dynamics Investigation (GEDI), and Gaofen (GF)-7. Among them, the ICESat-2/ATLAS implements the micro-pulse photon counting LiDAR technique, which is introduced in Section 2.4.3. The other systems use the full-waveform LiDAR. Here we mainly introduce the principles of spaceborne full-waveform LiDAR.

The spaceborne LiDAR, with high platform altitude and repeated, large-area observations, can provide reliable and multitemporal height measurement data, which have been widely used in mapping the surface vertical structure of the Earth and the surface topography of other planets. For example, the first LiDAR satellite ICESat/GLAS system detects the Earth's surface by emitting the laser beam, and each beam forms a nearly round footprint with a diameter of about 70 m on the ground (Schutz et al., 2005). The GEDI system, whose main mission is to monitor global forest biomass, carbon cycling, and biodiversity, has a footprint diameter of about 25 m on the ground (Dubayah et al., 2020). The GF-7 laser altimeter system has a footprint diameter of about 17 m.

The footprint diameter of spaceborne LiDAR is usually very large (>10 m) due to the high observation platform. The large footprint diameter might cause the ground vertical extent to overlap with the object vertical extent. The degree of overlapping is affected by the terrain slope, sensor observation orientation, and footprint diameter. Usually, to avoid signal aliasing as much as possible, the spaceborne LiDAR is designed with an on-nadir viewing mode, which means that the laser pulse is emitted vertically to the Earth's surface. Even so, sometimes, in order to collect more altimeter data, some spaceborne LiDAR systems are designed with multiple laser ground tracks by beam splitting and dithering, e.g., GEDI and

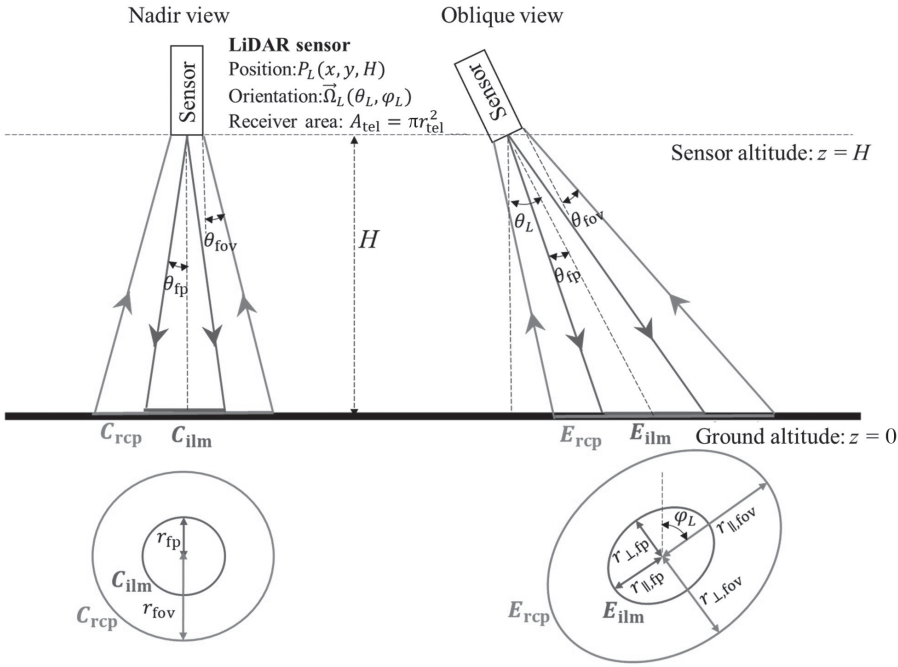

FIGURE 2.6 Pulse geometry of spaceborne LiDAR for nadir and oblique orientations, with the illumination and reception areas at the horizontal plane.

GF-7 systems. The modulation of multibeam splitting and dithering would cause the emitted laser pulse to be slightly incline with the milliradian level, which is usually negligible.

Figure 2.6 shows the geometry configuration of the laser pulses for nadir and oblique views. Generally, the LiDAR transmitter (laser) and receiver are considered to have the same location $P_L(x, y, H)$ and the same center orientation $\vec{\Omega}_L(\theta_L, \varphi_L)$. θ_L and φ_L are the zenith angle and azimuth angle of the LiDAR orientation, respectively; H is the sensor altitude. The laser can be regarded as a point source that emits laser power into a fixed conical solid angle with the divergence half-angle of θ_{fp}. The LiDAR receiver is usually a circular aperture with a radius of r_{tel} and the field of view (FOV) half-angle of θ_{fov}.

For the nadir view, the zenith and azimuth angles of the laser pulse are both 0°. A circular illumination area (i.e., C_{ilm} in Figure 2.6, also called the footprint) would be generated on the flat ground with an elevation of 0. The FOV area of the LiDAR receiver is also circular (i.e., C_{rcp} in Figure 2.6). The radii of the illumination area (r_{fp}) and FOV area (r_{fov}) are calculated as in Equation (2.27).

$$r_{fp} = H \cdot \tan\theta_{fp}, \quad r_{fov} = H \cdot \tan\theta_{fov} + r_{tel} \tag{2.27}$$

For the oblique view, the zenith angle and azimuth angle of the laser pulse are θ_L and φ_L, respectively. The illumination area (i.e., E_{ilm} in Figure 2.6) and the FOV

area (i.e., E_{rcp} in Figure 2.6) of the LiDAR system on flat ground are both elliptical. The long axis of the ellipse is in the same direction as the laser azimuth. In this case, the semi-major long axes (parallel to the laser azimuth) and the semi-minor axes (perpendicular to the laser azimuth) of the illumination area and the FOV area are calculated as in Equation (2.28).

$$\begin{aligned} r_{\perp,\text{fp}} &= \frac{H \cdot \tan\theta_{\text{fp}}}{\cos\theta_L}, \quad r_{\parallel,\text{fp}} = \frac{H \cdot \tan\theta_{\text{fp}}}{\cos^2\theta_L} \\ r_{\perp,\text{fov}} &= \frac{H \cdot \tan\theta_{\text{fov}}}{\cos\theta_L} + r_{\text{tel}}, \quad r_{\parallel,\text{fov}} = \frac{\dfrac{H \cdot \tan\theta_{\text{fov}}}{\cos\theta_L} + r_{\text{tel}}}{\cos\theta_L} \end{aligned} \tag{2.28}$$

where $r_{\perp,\text{fp}}$ and $r_{\parallel,\text{fp}}$ are the semi-major and semi-minor axes of the illumination area, respectively; $r_{\perp,\text{fov}}$ and $r_{\parallel,\text{fov}}$ are the semi-major and semi-minor axes of the FOV area, respectively.

2.3.2 Airborne LiDAR

Airborne LiDAR is defined as a LiDAR system equipped on the flying platforms such as drones and helicopters. It is also called airborne laser scanning (ALS), airborne laser altimetry (ALA), airborne laser topographic/terrain mapping (ALTM), and airborne laser mapping (ALM), among others. Although the names are slightly different, the principles are similar. That is, the LiDAR senor generates laser pulses and receives returned signals, and the scanning device controls the orientation of the LiDAR sensor. The main features of an airborne LiDAR system are as follows:

1. High precision. The vast majority of airborne LiDAR systems on the market use the pulsed laser ranging system with the consideration of its long measurement distance. However, it should be noted that the pulsed LiDAR requires advanced hardware and processing methods to achieve high ranging precision. Additionally, the ranging precision is influenced by the working environment to some extent. The LiDAR usually works best in dry, cold, and clear atmospheric conditions. In contrast, the LiDAR performs poor in the daytime with strong sunlight.
2. High power. The airborne LiDAR system scans the Earth's surface in the air. It requires a relatively high laser power, so that the returned laser energy through long-distance atmospheric loss and target absorption is strong enough to be recorded by the detector.
3. Small size. The aircraft's carrying capacity and volume are limited. It is required to load LiDAR equipment and carry operators in a limited space. Hence, airborne LiDAR systems are commonly designed to be small and lightweight.
4. Suitable wavelength. It would be better that the selected laser wavelength satisfies the following requirements: (1) the wavelength is located in the atmospheric window, so that the energy is less absorbed by the atmosphere;

(2) the wavelength is positioned where target has strong reflectance, so that the target can return a strong laser signal; (3) the wavelength is positioned where the detector sensitivity is high, so as to reduce the entry of light in other bands; and (4) it is safe for human eyes. Most existing airborne LiDAR systems usually use a 1064-nm, 1550-nm, 905-nm, or 532-nm wavelength. Some use other special wavelengths, but relatively few.

The airborne LiDAR systems scan the target in specific modes (Wehr & Lohr, 1999). The four main scanning modes are as follows:

1. Swing mirror scanning mode (Figure 2.7). This scanning mode drives the mirror with a motor to repeatedly swing within a certain angle, to achieve a series of scanning lines of laser beams on the ground. In one swing mirror cycle, the laser forms a periodic movement track on the ground, usually Z-shaped or sinusoidal. The advantage of the swing mirror mode is the simple principle, high scanning efficiency, and large scanning angle range. One of its significant disadvantages is that the laser spot density is not even, being denser at both ends and sparse in the middle, caused by the mechanical device continuously undergoing acceleration and deceleration processes

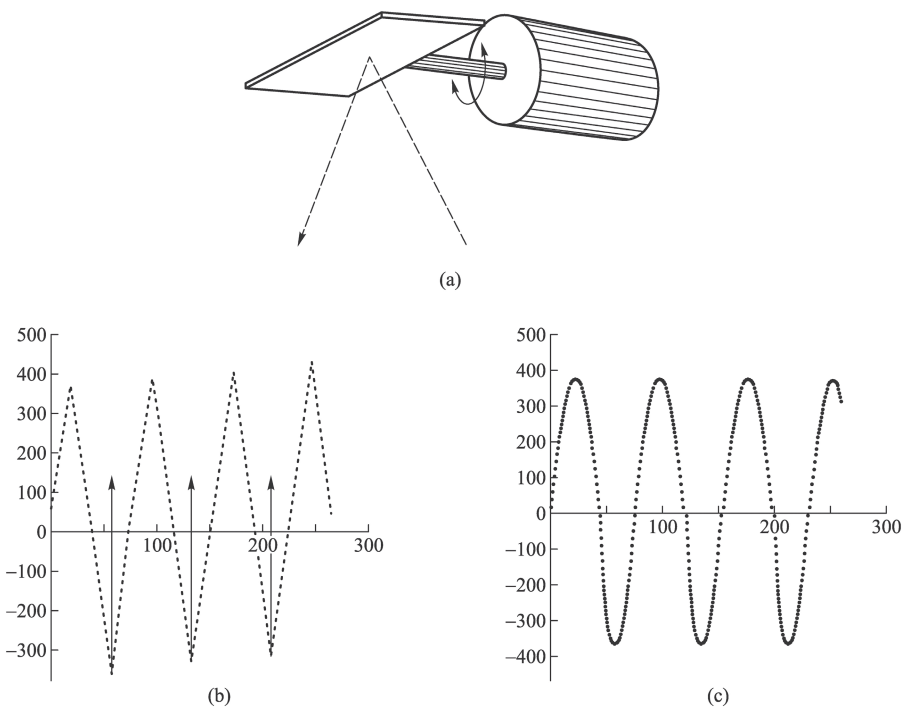

FIGURE 2.7 Swing mirror scanning mode of airborne LiDAR. (a) Scanning diagram; (b) Z-shaped ground scan line; (c) sinusoidal ground scan line.

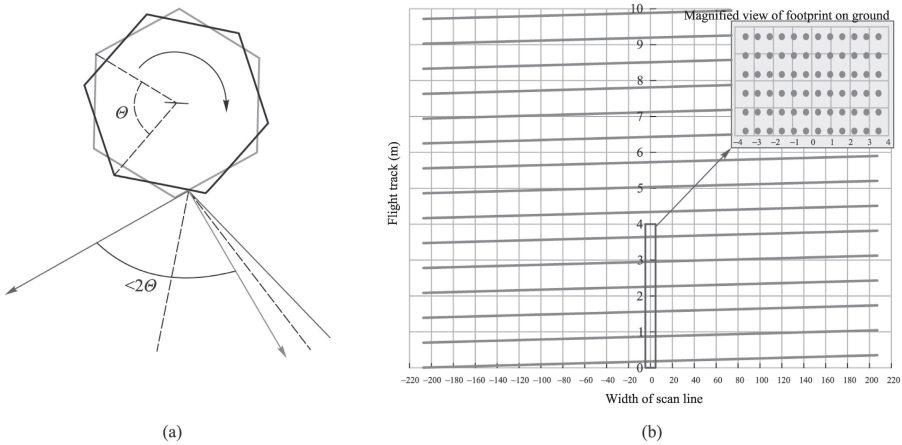

FIGURE 2.8 Rotating prism scanning mode of airborne LiDAR. (a) Scanning diagram; (b) parallel ground scan lines.

during the swing periods. Also, this mode has the disadvantages of poor seismic performance and high requirements for motor performance.

2. Rotate prism scanning mode (Figure 2.8). This mode causes the polygonal prism with the motor to rotate, causing the direction of the reflected laser beam to reciprocate within a certain range and scan on the ground. The scanning lines formed are usually parallel lines. Its advantage is that the obtained pulse density is uniform, as the angular velocity of the prism rotation is constant. The disadvantage is its relatively low scanning efficiency.
3. Elliptical scanning mode (Figure 2.9). This mode forms an elliptical scan line on the ground. With the movement of the sensor, a series of overlapping elliptical scan lines is presented on the ground. Its advantages include the simple scanning principle, good seismic performance, high ranging accuracy, and similar influence of inclination angle and atmospheric transfer on all laser points. However, the sampling points of the elliptical trajectory are

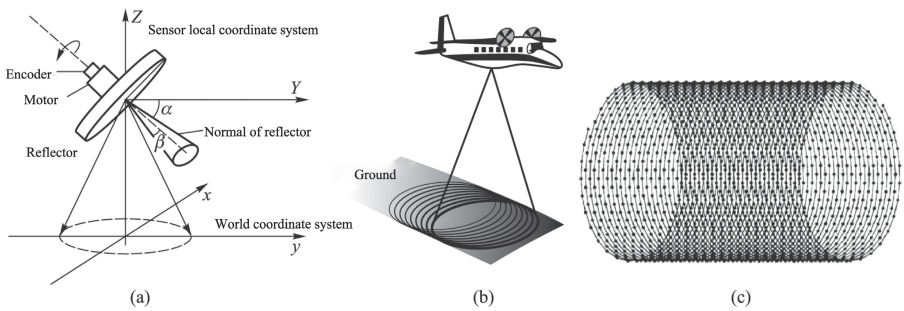

FIGURE 2.9 Elliptical scanning mode of airborne LiDAR. (a) Scanning diagram; (b) and (c) elliptical ground scan lines.

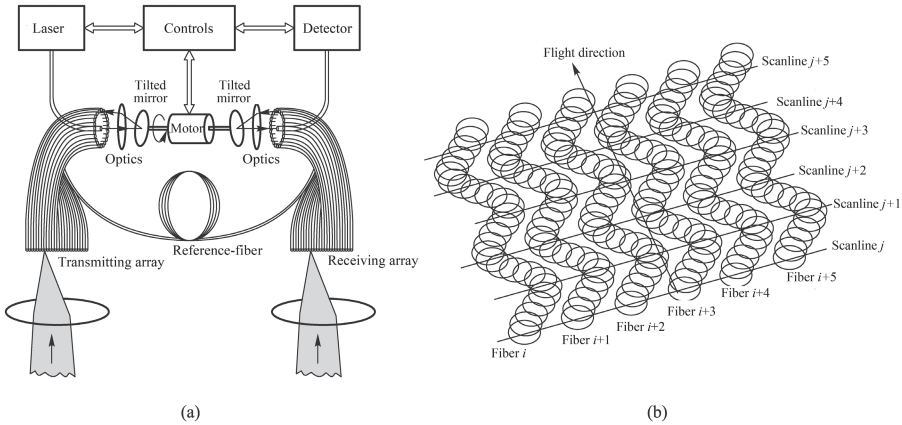

FIGURE 2.10 Fiber scanning mode of airborne LiDAR. (a) Scanning system; (b) ground scan lines.

unevenly distributed, which brings difficulties in terms of data processing. Also, the high repetition rate of the spots on the ground cause relatively low scanning efficiency.

4. Fiber scanning mode (Figure 2.10). Its principle is that the laser is transmitted through the fiber and enters the fiber group in turn; the terminals of the fiber group are arranged in a row; the fibers in different positions correspond to different fields of view; the lasers are emitted from the fiber arrays, in turn, to scan the target. The advantage is fast scanning speed and evenly distributed scanning points. The even points greatly shorten the data processing time.

2.3.3 Terrestrial LiDAR

The terrestrial LiDAR systems include fixed LiDAR and mobile LiDAR such as vehicle-mounted and backpack-based scanning systems. This section introduces the working principle of the terrestrial LiDAR system based on fixed tripod.

The terrestrial LiDAR fixes the 3D laser scanner on a tripod or other base station. The laser pulse is transmitted, then reflected by the object, and finally returns to the receiver. The distance between the scanner and object is calculated by the time delay of emitting and receiving the pulse. Also, the transverse scan angle θ and longitudinal scan angle φ of each pulse are recorded. Generally, the operator sets the transverse scan angle range $\Delta\theta$ and longitudinal scan angle range $\Delta\varphi$ to scan the scene in different directions. Also, the transverse scan angle resolution $\delta\theta$ and longitudinal scan angle resolution $\delta\varphi$ can be set to adjust the pulse density of terrestrial LiDAR. The number of emitted laser pulses (N_{tls}) can be calculated according to the scanning angle range and resolution, as in Equation (2.29).

$$N_{\text{tls}} = \frac{\Delta\theta}{\delta\theta} \cdot \frac{\Delta\varphi}{\delta\varphi} \tag{2.29}$$

Note the point cloud acquired by the terrestrial LiDAR is usually not uniform. Specifically, the objects close to the scanner have large point densities; otherwise, the point density is relatively small. Additionally, due to the occlusion between the objects, the single-station scanning cannot acquire comprehensive 3D information of all the objects in the scene. Therefore, in practice, the operators usually conduct multistation scanning with the terrestrial LiDAR and then splice the acquired data. More information about terrestrial laser scanning is provided in Section 3.2.

2.4 PRINCIPLES OF LiDAR WITH DIFFERENT METHODS OF DETECTION AND DIGITIZATION

Full waveform, discrete return, and photon counting are three common methods of detection and digitization for LiDAR systems (Mandlburger et al., 2019). Figure 2.11 shows how the three methods record the emitted and received signals when the laser penetrates a complex forest canopy. Among them, the full waveform LiDAR (fwLiDAR) and discrete return LiDAR (drLiDAR) record the returned signal through an analog-to-digital convertor, while the photon counting LiDAR (pcLiDAR) directly records the returned laser photons. For the two modes of analog electronic detection and direct detection, the received signal is composed of a valid laser signal and invalid noise signal. The noise signal is mainly caused by the sensor system and radiation from the sun. The difference is that the analog electronics detector converts the received laser power into the voltage signal, producing the returned intensity that varies over time. Also, the LiDAR with analog-to-digital record mode generally emits relatively high-power laser beams. The received valid laser signal has thousands of photons, which greatly exceeds the detector's internal noise and solar noise, i.e., high SNR. It helps to accurately identify a valid signal and

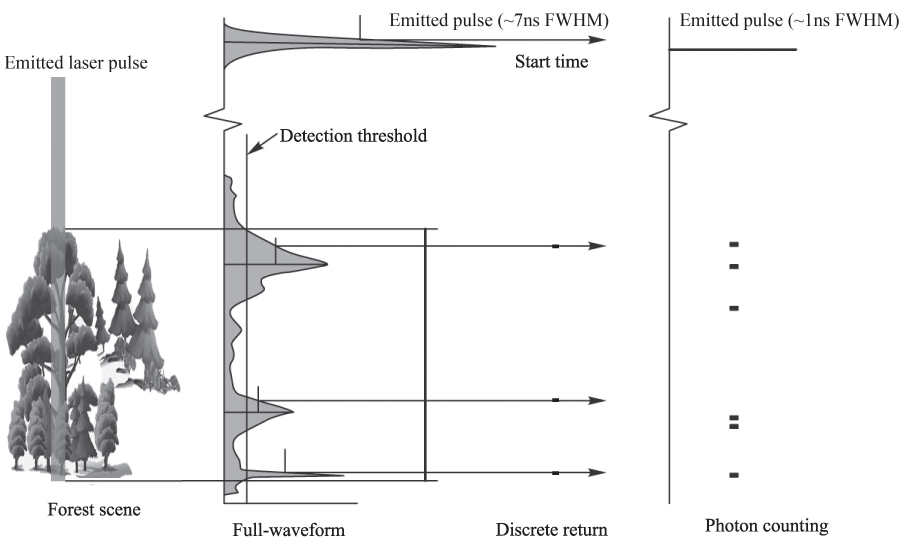

FIGURE 2.11 Principles of LiDAR with different methods of detection and digitization.

then calculate the distance between the sensor and target. In contrast, the photon counting mode employs a micro-pulse laser that emits a very weak laser beam. A single-photon detector is used to detect the returned photons and record their return time. However, the solar noise is usually obvious compared with the weak laser energy. This causes the low SNR of pcLiDAR.

We present the detailed principles of the fwLiDAR, drLiDAR, and pcLiDAR systems next.

2.4.1 Full-Waveform LiDAR

Full-waveform LiDAR uses an analog-to-digital converter (ADC) to digitize the returned laser power time series into a waveform. The waveform can characterize the complete vertical structure of the target, such as the gray shaded signal in Figure 2.11 (Harding et al., 2001). The pulse width of the returned waveform is influenced by the width of the emitted pulse and the vertical extent of the target within the footprint. For a target with a simple structure, such as a flat surface or building roofs, we would observe one returned laser peak. For a target with a complex structure, multiple return peaks are formed. For example, the laser hitting a building edge would be reflected by the building roof and the adjacent ground, thus forming two returned peaks. For objects with gaps, e.g., vegetation canopy, the laser pulse would irradiate the multiple objects in the transmitted direction, including leaves, stems, branches, and understory, forming a waveform signal with complex shapes. This process demonstrates the capability of LiDAR to capture vegetation's vertical structure.

Generally, the fwLiDAR system provides the digital recordings of emitted and received waveforms. The valid returned signal is calculated by detecting the first waveform location and the last waveform location exceeding the detected threshold. Then the ranging information is derived from the time delay between the center of the emitted pulse and the center of a valid received signal. The commonly used methods of calculating the pulse center include the centroid method, mean value method, and Gaussian fitting method (Tang et al., 2016).

Figure 2.12 displays the development of fwLiDAR systems. The earliest waveform sampling experiment for land terrain detection was the Airborne Oceanographic

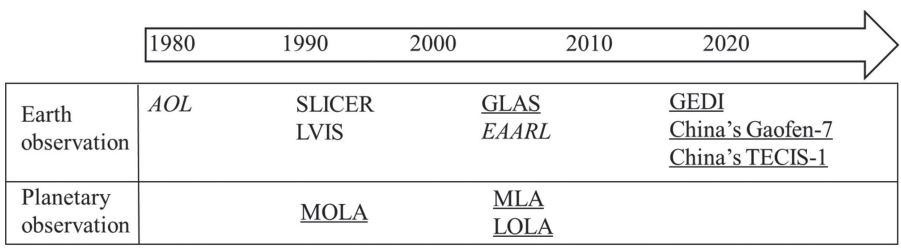

FIGURE 2.12 Existing full-waveform LiDAR systems.

LiDAR (AOL) system developed by the National Aeronautics and Space Administration (NASA) (Krabill et al., 1984). Subsequently, scientists developed multiple types of airborne fwLiDAR systems by combining the full-waveform LiDAR sensor with the airborne scanning device. The typical ones are Airborne Topographic Mapper (ATM) (Krabill et al., 2002), Scanning Lidar Imager of Canopies by Echo Recovery (SLICER) (Harding et al., 2000), Laser Vegetation Imaging Sensor (LVIS) (Blair et al., 1999), and Experimental Advanced Airborne Research LiDAR (EAARL) (Nayegandhi et al., 2005).

The main purpose of the SLICER and LVIS systems is to measure the vegetation canopy structure. They are designed with a large LiDAR footprint with a diameter larger than 10 m, of which obtained waveform sampling can reveal the forest vertical structure at a large scale. In contrast, the purpose of the ATM and EAARL systems is high-resolution topographic mapping with a small footprint (diameter <1 m). In the forest region, the small footprint cannot cover a tree crown completely. Hence, the acquired waveform cannot characterize the large-scale vegetation vertical structure.

In particular, the EAARL system uses a 532-nm laser and transmits the returned signal to channels with different ADC ratios. Its purpose is to adapt the significant variance of laser intensity returned from different scatter surfaces, such as land, vegetation, surface of the water, and bottom of the water. Hence, the EAARL system is capable of acquiring bathymetric data and land topographic data. If the returned intensity is not scaled properly, the recorded waveform signal would be saturated or distorted.

In addition to airborne platforms, full-waveform LiDAR is also equipped on spaceborne platforms to measure the Earth's 3D surface structure on a global scale. However, the existing LiDAR satellites equip the laser altimetry, rather than the laser scanner. The ICESat/GLAS is the world's first satellite-based, full-waveform laser altimeter for Earth observation. It was launched by NASA in 2003 and stopped working in 2009. In 2018, NASA carried the spaceborne full-waveform laser altimetry GEDI onto the ISS for global forest monitoring. China's GF-7, launched in 2019, and Terrestrial Ecosystem Carbon Inventory Satellite-1 (TECIS-1) launched in 2022, are both equipped with full-waveform laser altimeters. Their missions are global topography mapping and forest biomass monitoring, respectively. More details about the spaceborne LiDARs for Earth observation are introduced in Section 3.3 of this book.

Except for Earth observation, full-waveform LiDAR systems were also applied to other planetary surveying, such as the Mars Orbiter Laser Altimeter (MOLA) for Mars surface mapping (Smith et al., 1999), Messenger Laser Altimeter (MLA) for Mercury surface detection (Cavanaugh et al., 2007), and Lunar Orbiter Laser Altimeter (LOLA) for lunar surface detection (Smith et al., 2010).

2.4.2 Discrete Return LiDAR

In contrast to fwLiDAR, discrete return LiDAR records only a few locations in the waveform time series output by the detector. Threshold detection, recording the position exceeding the threshold in the waveform time series, is the simplest method of recording discrete return points. With this method, only one or a limited number of objects irradiated by the laser are detected, which greatly reduces the recorded data volume.

A crucial system parameter of the drLiDAR system is the maximum number of returns that can be recorded per pulse. The earliest drLiDAR can only record one point per pulse. The drLiDAR systems developed subsequently gradually record two, three, and up to five points per pulse. Taking as an example of Figure 2.11, the drLiDAR detector records three discrete points from the waveform time series. In addition to recording the 3D coordinates (X, Y, Z) of each discrete point, the intensity (I) of each discrete point is recorded by using the peak detection method and the Gaussian decomposition method. The peak detection method records the peak center of the waveform series as the point bin location and the peak amplitude as the point intensity. The Gaussian decomposition method performs the Gaussian decomposition algorithm on the waveform series. The center of the decomposed Gaussian component is recorded as the point bin location, and the amplitude is recorded as the point intensity. The time delay between the point bin location and the center location of the emitted pulse is used to calculate the distance between the target corresponding to each point and the sensor. Then combined with the global positioning system (GPS) and inertial measurement unit (IMU) data of the sensor, the 3D coordinates of each point are derived. Through multiple pulse laser scanning, massive discrete points with 3D geographic coordinates are generated, usually called a "point cloud".

The drLiDAR systems are usually designed with a small footprint and high pulse repetition frequency. Figure 2.13 shows an example of a discrete point cloud in a forest area acquired by a small-footprint drLiDAR system, recording up to four returns per pulse. The position, intensity, and corresponding return number of each discrete point depend on the spatial structure and optical properties of the surface objects.

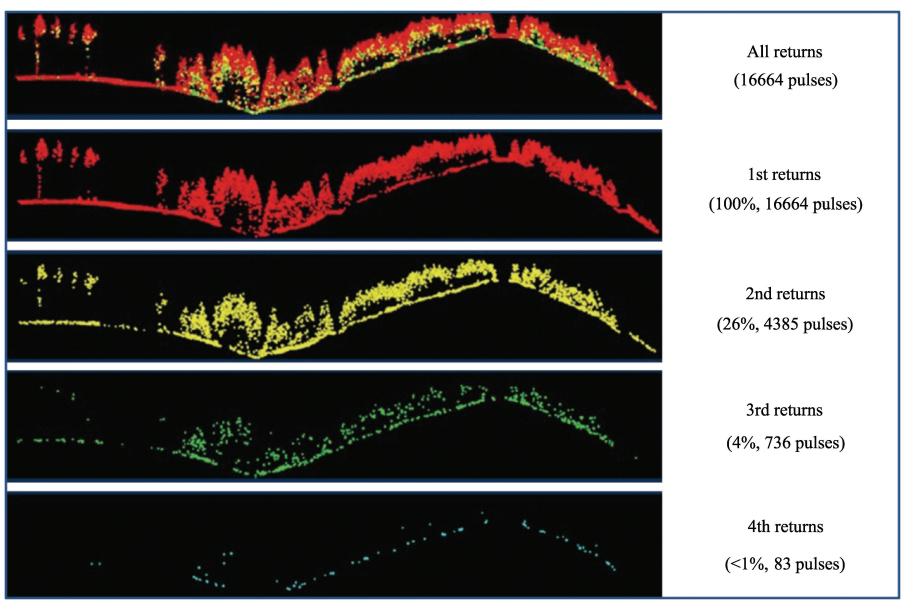

FIGURE 2.13 Multiple returns of discrete return LiDAR.

According to the LiDAR equation (Equation 2.20), the return laser intensity is related to the intercepted area and reflectance of the corresponding object. Once the object property makes it generate a strong enough laser intensity to exceed the detection threshold, the detector records it as one discrete point. The first discrete point corresponds to the location of the waveform signal that exceeds the intensity threshold for the first time, which might come from the top of the vegetation canopy, or the interior of the vegetation canopy, or the ground. The last discrete point corresponds to the location of the detected waveform signal for the last time, which might come from the interior of the vegetation canopy or from the ground. Sometimes, some ground points can be detected in the vegetation-covered region. This is because the laser pulse can penetrate gaps in the vegetation canopy to reach the ground. The frequency of detecting the ground under the canopy depends on many factors, such as canopy gap fraction, laser scanning angle, beam divergence angle, footprint diameter, pulse repetition frequency, and ground reflectance.

2.4.3 Photon Counting LiDAR

Different from the first two detection modes, the photon counting LiDAR records the arrival time of each laser photon to achieve the sampling of the target vertical structure (Figure 2.11). Through detection of multiple pulses, the accumulated photon points potentially construct the spatial structure of the Earth's surface (Howland et al., 2011).

Full-waveform LiDAR and discrete return LiDAR systems usually emit high-intensity laser pulses, thus obtaining the laser signal with a high SNR. In contrast, photon counting systems emit very weak micro-pulses, with only a small number of photons returning per pulse. Hence, a single-photon detector is required to detect the returned photons. The very narrow pulse width of the emitted pulse (<1 ns), low jitter of the detector, and high resolution of the timing electronics make the pcLiDAR reach decimeter-level ranging accuracy. Note that the single photon detector of pcLiDAR usually has some dark counts, i.e., system noise. The dark counts are quantified by the parameter "dark count rate," which is the detector's thermal noise count rate without any light source. The dark count rate is usually very low compared with the frequency of received photon counts.

In addition to good detection sensitivity and low dark count noise, the single photon detector can detect multiple photons for one pulse. That is, a few photon points are recorded for each pulse. This characteristic is determined by a key detector parameter, dead time, which is defined as the period needed by the detector to recover from a detection event, before another photon can be detected. The dead time of a single photon detector is very short, so that the pcLiDAR can distinguish between consecutive photons returned from closely spaced surfaces, e.g., the returns from vegetation canopy. This characteristic of pcLiDAR not only can detect multiple photons for one pulse, improving the photon density, but also record the acquired data in the form of discrete photon points.

When the photon counting LiDAR system works in the daytime, the solar radiations reflected from the Earth's surface and clouds are the main noise source. The solar noise rate is influenced by the surface reflectance, solar irradiance, atmospheric

environment, and sensor configuration, among other things. Generally, there is a filter with a very narrow bandwidth located before the pcLiDAR detector. Its purpose is to prevent the electromagnetic wave other than the laser band from entering the detector. Also, pcLiDAR systems are typically designed to have a small FOV, confining the received signal within the FOV region.

To summarize, the pcLiDAR has the characteristics of low laser power, small receiving FOV, and high pulse repetition frequency. Compared with the analog-to-digital LiDAR, it has strong potential in measuring fine 3D structures of the Earth's surface at the large scale. Also, the output of pcLiDAR is the discrete photon points with specific 3D spatial coordinates. This makes the existing point cloud algorithms of data processing, analysis and visualization, and the software potentially suitable for pcLiDAR data.

2.5 INFLUENCE MECHANISM OF THE SKY ON THE LiDAR SIGNAL

The LiDAR is inevitably affected by the surrounding environment during the measurement process, especially spaceborne and airborne LiDAR. Figure 2.14 shows the radiative transfer process of sunlight energy (red arrows) and laser energy (green

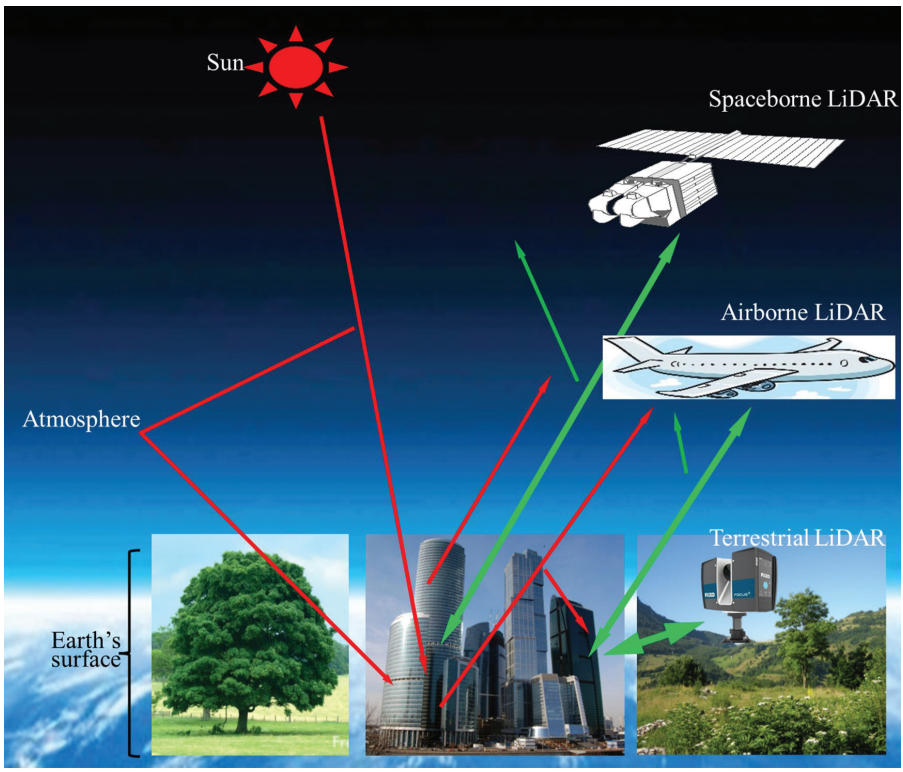

FIGURE 2.14 Radiative transfer process of sunlight and laser among the atmosphere, Earth's surface, and LiDAR sensors.

arrows) among the atmosphere, Earth's surface, and sensors. It is seen that in the transmitted process, the laser pulse interacts with not only the Earth's surface but also the atmosphere layer. Also, the sensor receives some noise signals from the sunlight, including the solar radiation scattered by the atmosphere and surface. This section presents the influential mechanisms of the atmosphere and sunlight on the LiDAR signal.

2.5.1 Effect of Atmosphere

The LiDAR systems usually use a laser with a short wavelength. The atmosphere will absorb and scatter the laser signal, causing the LiDAR sensor to be very sensitive to the atmosphere (Qiang et al., 2000). In particular, rain, dust, fog, and haze in the atmosphere interfere with the laser signal. The main influences of atmosphere on the LiDAR signal include (1) energy attenuation: the atmospheric gas, aerosol, dust, fog, and rain would absorb or scatter the laser signal, causing the attenuation of laser energy, and (2) laser refraction: the uneven spatial distribution of the atmosphere causes the laser to be refracted rather than directly fired. Additionally, atmospheric turbulence affects the fluctuation of laser energy distribution on the beam cross section and the beam expansion and drift. Atmospheric absorption in the laser transmission path causes variations of the density gradient and refractive index of air, i.e., atmospheric thermal halo effect. This results in the nonlinear thermal distortion of the laser beam. The effects of atmospheric attenuation and atmospheric refraction in the LiDAR transmission process are introduced in detail next.

2.5.1.1 Atmospheric Attenuation

A laser has excellent monochromaticity. With the assumption that the atmosphere is uniform or stratified uniform, the laser energy through the atmosphere layer is expressed by the Beer-Lambert law:

$$I(\lambda, z) = I(\lambda, 0)\mathrm{e}^{-\sigma(\lambda)\cdot z} \tag{2.30}$$

where $I(\lambda,0)$ and $I(\lambda,z)$ are the initial laser intensity and the laser intensity passing through the atmosphere with a thickness of z, respectively, λ is the laser wavelength, and $\sigma(\lambda)$ is the atmospheric attenuation coefficient related to the wavelength.

The atmospheric attenuation includes two parts: atmosphere absorption and atmosphere scattering. Therefore, $\sigma(\lambda)$ can be expressed as Equation (2.31):

$$\sigma\left(\lambda\right) = \sigma_{\mathrm{m}} + K_{\mathrm{m}} + \sigma_{\mathrm{a}} + K_{\mathrm{a}} \tag{2.31}$$

where σ_{m} is the gas molecular scattering coefficient, K_{m} is the gas molecular absorption coefficient, σ_{a} is the aerosol scattering coefficient, and K_{a} is the aerosol absorption coefficient.

The absorption coefficient of a gas molecule and aerosol for laser light can be calculated based on the transmission theory of monochromatic light. When the laser pulse is transmitted in the lower atmosphere below 20 km, the width of the atmosphere absorption spectrum is mainly determined by the pressure broadening

caused by gas molecular collision. The gas molecular absorption coefficient of the low atmosphere layer (K_{mL}) is expressed as Equation (2.32):

$$K_{\mathrm{mL}} = \frac{S}{\pi} \cdot \frac{\gamma_L}{(v - v_0)^2 + {\gamma_L}^2} \tag{2.32}$$

where v is light frequency, v_0 is the frequency of the center of the laser spectrum, S is the integrated intensity of the laser spectrum, γ_L is a half-width of the Lorentz spectral line, and S and γ_L are both influenced by atmospheric temperature and pressure.

In the upper atmosphere above 60 km, the gas molecular absorption coefficient (K_{mD}) is calculated by the Doppler-broadened line-shaped function, as in Equation (2.33).

$$K_{\mathrm{mD}} = \frac{S}{\gamma_D} \cdot \left(\frac{\ln 2}{\pi}\right)^{1/2} \cdot \mathrm{e}^{-\ln 2 \cdot (v - v_0)^2 / {\gamma_D}^2} \tag{2.33}$$

where γ_D is the Doppler line half-width. In the middle atmosphere at altitudes of 20–60 km, the collision broadening and Doppler broadening exist simultaneously. The gas molecular absorption coefficient (K_{mV}) is given by the Voigt line shape, as in Equation (2.34).

$$K_{\mathrm{mV}} = \frac{S \cdot {\gamma_L}^2}{\gamma_D \cdot \sqrt{\pi}} \cdot \int_{-\infty}^{\infty} \mathrm{e}^{-t^2} \cdot \left[\left(\frac{\gamma_L}{\gamma_D}\right)^2 + \left(\frac{v_0 - v}{\gamma_D} - t\right)^2\right] \mathrm{d}t \tag{2.34}$$

The atmospheric scattering effect includes Rayleigh scattering by atmosphere gas molecular and Mie scattering by spherical or quasi-spherical particles such as raindrops, fog droplets, haze, and aerosols.

Rayleigh scattering occurs when the laser wavelength is much larger than the particle size, and the gas molecular scattering coefficient σ_{m} is expressed as in Equation (2.35).

$$\sigma_{\mathrm{m}} = \frac{8\pi^2}{3} \cdot \frac{(n^2 - 1)^2}{N_s^2 \lambda^4} \cdot \frac{6 + 3\delta_p}{6 - 7\delta_p} \tag{2.35}$$

where n is the particle refractive index, N_s is particle density in the air, and δ_p is the depolarization factor. Based on experience, σ_{m} is expressed as:

$$\sigma_{\mathrm{m}} = 2.677 \times 10^{-17} P v^4 / T \tag{2.36}$$

where P is pressure and T is temperature.

Mie scattering occurs when the size of atmospheric particles is comparable to the laser wavelength. Generally speaking, the Mie scattering is the aerosol scattering. The total atmosphere attenuation coefficient can be approximately expressed as:

$$\sigma_{\mathrm{T}} = \sigma_{\mathrm{a}} \cdot \sigma_{\mathrm{m}} = \frac{3.912}{V_{\mathrm{m}}} \cdot \left(\frac{0.55}{\lambda}\right)^b \tag{2.37}$$

where V_{m} is visibility (unit: km), which is defined as the maximum distance at which the human eye can distinguish the target, and the coefficient b is related to the visibility. For normal visibility (6 km ≤ V_{m} ≤ 20 km), $b = 1.3$; For good visibility ($V_{\mathrm{m}} > 20$ km), $b = 1.6$; For poor visibility ($V_{\mathrm{m}} < 6$ km), $b = 0.585 \cdot V_m^{1/3}$.

In general, the scattering coefficients of atmospheric gas molecules and aerosols obey the negative exponential distribution law of altitude. That is, with the increase of atmospheric altitude, the scattering coefficients decrease rapidly. For example, the aerosol scattering coefficient above 5 km decreases more than an order of magnitude compared with that on the ground. Additionally, in actual situations, the aerosol scattering is the main interaction process in the low atmosphere layer near the ground, while the contributions of gas scattering and aerosol scattering are comparable in the middle and upper atmosphere layers.

2.5.1.2 Atmospheric Refraction

The atmospheric refraction effect is defined as the phenomenon that when the laser passes through the atmosphere, the light transmission path is bent and the traveled distance increases, caused by uneven atmosphere density or existence of an atmospheric refractive index gradient. Its effect on LiDAR measurement is to increase the measured error of target bearing and distance.

The atmosphere density varies with the altitude, so the atmosphere refractive index for light at different altitudes is different. The refractive index n is related to the laser wavelength λ, atmosphere temperature T, humidity e, and pressure P, which is generally expressed as Equation (2.38):

$$n = 1 + N\left(\lambda, T, P, e\right) \tag{2.38}$$

where N is the refractive index modulus with the unit of order 10^{-6}. For the standard atmosphere ($P = 1$ atm, $T = 288.15$ K, $e = 0$), the atmosphere refractive index from visible to near-infrared band is expressed as Equation (2.39):

$$N_0(\lambda) = 272.5794 + 1.5932\lambda_0^{-2} + 0.015\lambda_0^{-4} \tag{2.39}$$

where $N_0(\lambda)$ is the standard atmospheric refractive index in the wavelength λ.

The atmospheric refractive index $N(\lambda)$ in any atmospheric condition can be calculated from the standard atmospheric refractive index $N_0(\lambda)$, as in Equation (2.40):

$$N(\lambda) = N_0(\lambda) \times \left(2.8434 \times 10^{-3} \times \frac{P}{T} - 0.1127 \times \frac{e}{T} \right) \tag{2.40}$$

The atmosphere attenuation effect in the LiDAR radiative transfer process was introduced earlier. For practical application, other complex effects, such as atmospheric turbulence and thermal halo, should also be considered. The propagation of the LiDAR signal in the atmosphere directly affects the sensor performance. Therefore, it is necessary to fully understand the atmospheric propagation characteristics of the laser, find ways to avoid or reduce the atmospheric effects, and select the suitable working wavelength and working mode for the LiDAR system according to

the laser propagation mechanism in the atmosphere to finally achieve the high atmosphere transmittance of the laser beam.

2.5.2 Effect of Sunlight

The existence of sunlight makes the LiDAR RS radiative transfer process a RS configuration of "two light sources (laser and sunlight) + one sensor" (Figure 2.15). The laser source is regarded as a point light source emitting laser power onto a fixed conical solid angle. The emitted laser only lasts for a short time. The sun emits solar energy from an infinite distance to the Earth's surface. Without the consideration of atmospheric scattering, the sunlight reaching the Earth's surface can be regarded as continuous parallel beams with the same direction. Once the sunlight irradiates the surface within the LiDAR receiving FOV, the solar energy might be scattered by the surface and then received by the LiDAR sensor. The influence of sunlight on the LiDAR signal is different for different types of LiDAR systems. Next we introduce its influences separately.

The large-footprint LiDAR systems (footprint diameter >10 m, usually mounted on spaceborne and airborne platforms) have a relatively large laser illumination region and receiving FOV region. The solar energy scattered into the LiDAR sensor

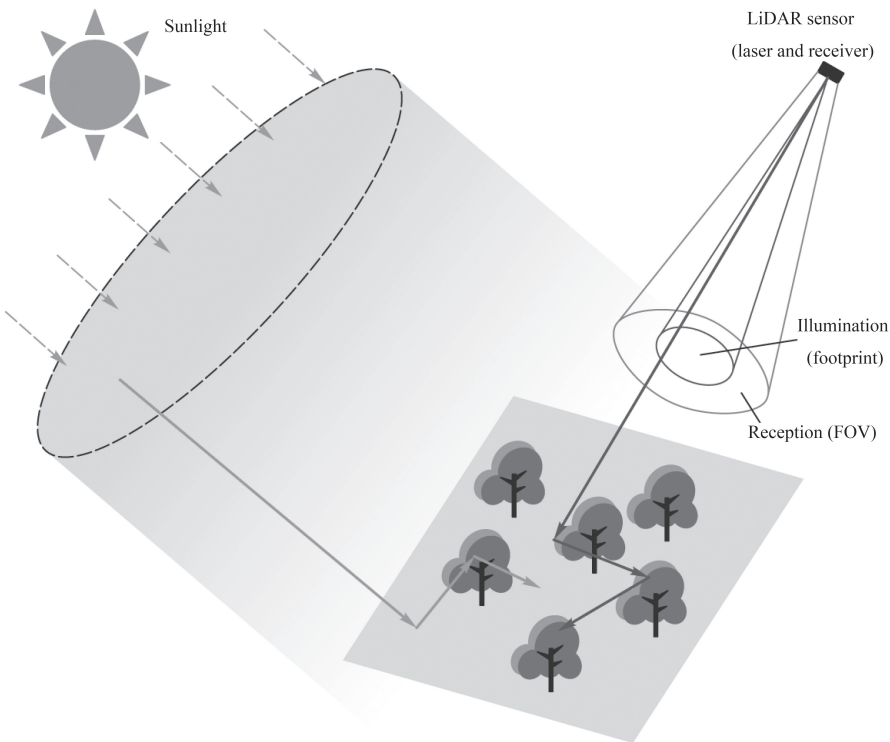

FIGURE 2.15 Radiative transfer process of dual light sources (laser and sunlight).

is relatively high. In contrast, the small-footprint LiDAR systems (footprint diameter <10 m, usually equipped on spaceborne and airborne platforms) have a small illumination region and receiving FOV region. Only very little solar energy enters the LiDAR receiver. Hence, the sunlight has less effect on the small-footprint LiDAR signal.

For fwLiDAR, the influence of sunlight is shown as a relatively constant solar background noise. This is because the persistence of sunlight makes its radiation transfer process to adopt a dynamic equilibrium state, and the solar energy entering the LiDAR receiver almost remains constant. The solar noise of a LiDAR waveform can be easily removed by setting a background noise threshold.

For drLiDAR, the solar noise is extremely weak due to the small footprint. Also, drLiDAR usually records discrete point clouds from waveform data by using the threshold detection method, peak detection method, or Gaussian decomposition method. The constant and very weak solar noise does not affect the calculated 3D coordinates of these points.

For pcLiDAR, the detector treats solar photons in the filter band the same as laser photons. That is, the recorded photons are a mixture of solar photons, laser photons, and dark noise photons. Due to the extremely low laser energy of pcLiDAR, solar noise has a significant effect on the pcLiDAR data acquired in the daytime. Furthermore, due to the persistence of solar irradiation, the solar noise photons exist over the vertical extent of the entire scene, greatly increasing the density of the photon points. Therefore, in order to obtain accurate surface structural information, the solar noise photons must be removed by postprocessing algorithms. The denoising algorithms for photon counting data are introduced in Section 4.3.

2.6 SUMMARY

This chapter first introduces the ranging principle of LiDAR systems, including pulsed LiDAR and phased LiDAR. Then, the LiDAR radiation principle is introduced based on the LiDAR equation. The commonly used LiDAR waveform model and LiDAR radiative transfer model are also presented. After that, we systematically introduce the principles of LiDAR systems with different platforms (spaceborne/airborne/terrestrial) and different methods of detection and digitization (full waveform/discrete return/photon counting). Finally, the influence mechanisms of the sky environment (atmosphere and sunlight) on the LiDAR signal are analyzed.

EXERCISES

1. Why do most LiDAR systems employ the pulse ranging method?
2. What is the ranging principle of the pulsed LiDAR? How is it different from that of the phased LiDAR?
3. Please write the LiDAR equation and explain the meaning of each parameter.
4. Please list three LiDAR 3D radiation transfer models and introduce their core ideas.
5. Please state the commonly used scanning modes of airborne LiDAR systems and their advantages and disadvantages.

6. Please explain the differences of the LiDAR systems with three different methods of detection and digitization, i.e., full waveform LiDAR, discrete return LiDAR, and photon counting LiDAR.
7. How are the sizes of the LiDAR footprint and field of view (FOV) calculated?

REFERENCES

Blair, J. B., & Hofton, M. A. (1999). Modeling laser altimeter return waveforms over complex vegetation using high resolution elevation data. *Geophysical Research Letters*, *26*(16), 2509–2512.

Blair, J. B., Rabine, D. L., & Hofton, M. A. (1999). The laser vegetation imaging sensor: A medium-altitude, digitization-only, airborne laser altimeter for mapping vegetation and topography. *ISRPS Journal of Photogrammetry and Remote Sensing*, *54*(2–3), 115–122.

Cavanaugh, J. F., Smith, J. C., Sun, X. L., Bartels, A. E., Ramos-Izquierdo, L., Krebs, D. J., McGarry, J. F., Trunzo, R., Novo-Gradac, A. M., Britt, J. L., Karsh, J., Katz, R. B., Lukemire, A. T., Szymkiewicz, R., Berry, D. L., Swinski, J. P., Neumann, G. A., Zuber, M. T., & Smith, D. E. (2007). The mercury laser altimeter instrument for the MESSENGER mission. *Space Science Reviews*, *131*(1–4), 451–479.

Dai, Y. (2002). *The principle of LiDAR* (in Chinese). National Defense Industry Press.

Disney, M. I., Lewis, P. E., Bouvet, M., Preto-Blanco, A., & Hancock, S. (2009). Quantifying surface reflectivity for spaceborne LiDAR via two independent methods. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(9), 3262–3271.

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P. L., Qi, W. L., & Silva, C. (2020). The global ecosystem dynamics investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, *1*, 100002.

Gastellu-Etchegorry, J.-P., Yin, T., Lauret, N., Grau, E., Rubio, J., Cook, B. D., Morton, D. C., & Sun, G. (2016). Simulation of satellite, airborne and terrestrial LiDAR with DART(i): Waveform simulation with quasi-Monte Carlo ray tracing. *Remote Sensing of Environment*, *184*, 418–435.

Goodenough, A. A., & Brown, S. D. (2017). DIRSIG5: Next-generation remote sensing data and image simulation framework. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(11), 4818–4833.

Govaerts, Y. M., & Verstraete, M. M. (1998). Raytran: A Monte Carlo ray-tracing model to compute light scattering in three-dimensional heterogeneous media. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(2), 493–505.

Harding, D. J., Blair, J. B., & Rabine, D. L. (2000). *SLICER airborne laser altimeter characterization of canopy structure and sub-canopy topography for the BAREAS northern and southern study regions: Instrument and data product description*. National Aeronautics and Space Administration, Goddard Space Flight Center.

Harding, D. J., Lefsky, M. A., Parker, G. G., & Blair, J. B. (2001). Laser altimeter canopy height profiles: Methods and validation for closed-canopy, broadleaf forests. *Remote Sensing of Environment*, *76*(3), 283–297.

Howland, G. A., Dixon, P. B., & Howell, J. C. (2011). Photon-counting compressive sensing laser radar for 3D imaging. *Applied Optics*, *50*(31), 5917–5920.

Huang, H., & Wynne, R. H. (2013). Simulation of LiDAR waveforms with a time-dependent radiosity algorithm. *Canadian Journal of Remote Sensing*, *39*, S126–S138.

Kobayashi, H., & Iwabuchi, H. (2008). A coupled 1-D atmosphere and 3-d canopy radiative transfer model for canopy reflectance, light environment, and photosynthesis simulation in a heterogeneous landscape. *Remote Sensing of Environment*, *112*(1), 173–185.

Krabill, W. B., Abdalati, W., Frederick, E. B., Manizade, S. S., Martin, C. F., Sonntag, J. G., Swift, R. N., Thomas, R. H., & Yungel, J. G. (2002). Aircraft laser altimetry measurement of elevation changes of the Greenland ice sheet: Technique and accuracy assessment. *Journal of Geodynamics*, *34*(3–4), 357–376.

Krabill, W. B., Collins, J., Link, L., Swift, R., & Butler, M. (1984). Airborne laser topographic mapping results. *Photogrammetric Engineering and Remote Sensing*, *50*(6), 685–694.

Lai, X. (2010). *Basic principles and applications of airborne LiDAR* (in Chinese). Publishing House of Electronics Industry.

Mandlburger, G., Lehner, H., & Pferfer, N. (2019). A comparison of single photon and full-waveform LiDAR. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *4*, 397–404.

Nayegandhi, A., Brock, J. C., & Wright, C. W. (2005). Classifying vegetation using NASA's Experimental Advanced Airborne Research LiDAR (EAARL) at Assateague island national seashore. American Society for Photogrammetry and Remote Sensing - Annual Conference 2005, 2, 769–777.

Ni-Meister, W., Jupp, D. L. B., & Dubayah, R. (2001). Modeling LiDAR waveforms in heterogeneous and discrete canopies. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(9), 1943–1958.

North, P. R. J., Rosette, J. A. B., Suarez, J. C., & Los, S. O. (2010). A Monte Carlo radiative transfer model of satellite waveform LiDAR. *International Journal of Remote Sensing*, *31*(5), 1343–1358.

Pharr, M., Jakob, W., & Humphreys, G. (2016). *Physically based rendering: From theory to implementation*. Morgan Kaufmann.

Qiang, X., Zhang, H., Tu, Q., Yuan, R., & Li, Z. (2000). Atmospheric attenuation effects of Ladar signals (in Chinese). *Journal of Applied Optics*, *21*(4), 21–25.

Schutz, B. E., Zwally, H. J., Shuman, C. A., Hancock, D., & DiMarzio, J. P. (2005). Overview of the ICESat mission. *Geophysical Research Letters*, *32*(21), 97–116.

Smith, D. E., Zuber, M. T., Neumann, G. A., Lemoine, F. G., Mazarico, E., Torrence, M. H., McGarry, J. F., Rowlands, D. D., Head, J. W., Duxbury, T. H., Aharonson, O., Lucey, P. G., Robinson, M. S., Barnouin, O. S., Cavanaugh, J. F., Sun, X. L., Liiva, P., Mao, D.-D., Smith, J. C., & Bartels, A. E. (2010). Initial observations from the lunar orbiter laser altimeter (LOLA). *Geophysical Research Letters*, *37*(18), L18204.

Smith, D. E., Zuber, M. T., Solomon, S. C., Phillips, R. J., Head, J. W., Garvin, J. B., Banerdt, W. B., Muhleman, D. O., Pettengill, G. H., Neumann, G. A., Lemoine, F. G., Abshire, J. B., Aharonson, O., Brown, C. D., Hauck, S. A., Ivanov, A. B., Mcgovern, P. J., Zwally, H. J., & Duxbury, T. C. (1999). The global topography of Mars and implications for surface evolution. *Science*, *284*(5419), 1495–1503.

Sun, G., & Ranson, K. J. (2000). Modeling LiDAR returns from forest canopies. *IEEE Transactions on Geoscience and Remote Sensing*, *38*(6), 2617–2626.

Tang, X., Li, G., Gao, X., & Chen, J. (2016). The rigorous geometric model of satellite laser altimeter and preliminarily accuracy validation (in Chinese). *Acta Geodaetica et Cartographica Sinica*, *45*(10), 1182–1191.

Wagner, W., Ullrich, A., Ducic, V., Melzer, T., & Studnicka, N. (2006). Gaussian decomposition and calibration of a novel small-footprint digitizing airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *60*(2), 100–112.

Wehr, A., & Lohr, U. (1999). Airborne laser scanning-an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, *54*(2–3), 68–82.

Yang, X., Wang, Y., Yin, T., Wang, C., Lauret, N., Regaieg, O., Xi, X., & Gastellu-Etchegorry, J.-P. (2022). Comprehensive LiDAR simulation with efficient physically-based DART-Lux model (I): Theory, novelty, and consistency validation. *Remote Sensing of Environment*, *272*, 112952.