White Paper Title

Harvey Weinberg, MicroTech Ventures and Pier-Olivier Hamel, indie Semiconductor

Introduction

As LIDAR has become a hot topic in the sensor world, mostly thanks to efforts in the ADAS and autonomous driving sector, a debate has emerged as to whether direct-detection (or time-of-flight) or coherent (Frequency Modulated Continuous Wave, for example) photon detection is best. In truth, "best" depends very much on the application. LIDAR is used in a wide variety of applications from traffic management, to driver assistance and autonomous driving, to ground mapping, to meteorological applications. It should be no surprise that the importance of different LIDAR performance metrics – maximum range, accuracy, interference immunity, cost, etc. – vary from application to application. Even within the same application, certain system choices may skew the importance of one parameter or another. This paper aims to discuss the different characteristics of direct and coherent detection in order to educate those interested in LIDAR and allow them to make informed system choices.

Example LiDAR point cloud

Historical Background

LIDAR was first conceptualized in the 1930's – around the same time as radar was developed. However, it wasn't until the early 1960's, when the first lasers were developed, that LIDAR became a reality. Coherent Frequency Modulated Continuous Wave (FMCW) radar had been developed in the mid-1930's[1] and shortly thereafter researchers set to work to bring the benefits of coherent detection to light-based ranging. During the 1960s, multiple researchers were demonstrating early FMCW LiDAR systems. Since then, LIDAR has found use in dozens of applications, each one with its unique list of challenges that developers had to overcome. The rise of optical telecommunications gave another boost to LIDAR from developments in advanced lasers and improved modulation techniques funded by the huge number of research dollars poured into optical telecommunications. Just as with radar, early optical telecommunications relied on pulsed, direct detection techniques. By 2008, coherent detection started to take over[2]. Today the industry is fortunate to have the benefit of over half a century of development in hardware and signal processing advances that were developed for radar and optical communications to draw on in bringing the "next" LIDAR system to life.

Photon Detection

At its most basic, a LIDAR system emits photons and calculates how long those photons took to reach a target and return. While there are many aspects of a LIDAR system to consider (what wavelength to use, scanning method, how to deal with interference, etc.), the choice of how we detect those returning photons drives almost every other system choice. There are, essentially, two methods of photon detection:

Direct detection - A laser pulse is fired which starts a timer. The timer is stopped when the echo of the laser pulse is received. We do not consider the phase of the photons, simply their existence (amplitude) and return timing. As the speed of light is known and invariant, we calculate the distance to the target as $\frac{\Delta t}{2}c$ where Δt is the time between the start of photon transmission and the leading edge of photon reception (as shown in Figure 1).

Coherent detection - A modulated laser is on for a longer period of time and the return signal is optically mixed with a sample of the transmitted photodetection (called the local oscillator) before photodiodedetection. This optical mixing results in the receive signal being amplified by the local oscillator. By using a sample of the transmit signal we are assured that the phase relationship between transmit and receive channels is preserved (or coherent). As with direct detection, distance is calculated by measuring the time between photon transmission and reception. But in the case of coherent detection, modulation is applied to the continuously (or quasi-continuously) transmitted signal. As the laser is transmitting continuously the echo timing is determined by appropriate demodulation, which requires more signal processing than direct detection. With coherent detection we can directly measure velocity instantaneously (not by measuring target movement over multiple frames as one would do with direct detection) by detecting the frequency shift of the returned signal caused by Doppler.

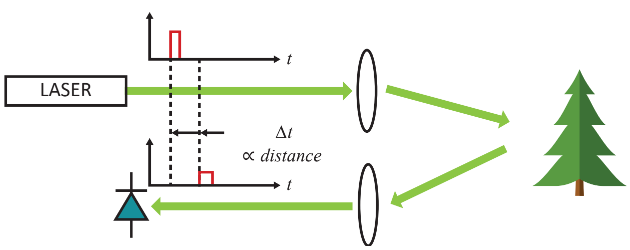


Figure 1. Graphic description of a direct detection LIDAR system

We will begin by describing direct detection in detail as it is conceptually simpler.

Direct Detection

Direct detection systems employ a pulsed laser to emit short (a few ns) bursts of light. The LIDAR sensor then measures the time needed to receive the reflected light pulse. By analyzing the time it takes for the light to travel to the target and back, it calculates the distance to objects in the environment.

Direct detection is suitable, if only modest performance (say, under 50m range) is needed. There is no need for a tunable, single mode laser (beam quality or coherence length, for example, are not critical) as it's simply a source of lots of photons squeezed into a short time period. The laser driver circuit is simplified as there is no need to modulate the laser; the modulator's task is to inject a lot of inject a lot of current into the laser within a few ns. The precision requirements for optics are relaxed as there is little concern about wavefront distortion.

Mathematically the return power in a direct detection-based system can be expressed as;

- $$Return \propto Power_{TX} \times \frac{Target\ Cross\ Section}{Illuminated\ Area} \times \frac{Receive\ Area}{\pi \cdot Range^2} \qquad \text{Eq. 1}$$

As one would intuitively suspect, we see that the return power drops as the square of the range. Likewise, return power also diminishes linearly as the illuminated area grows. Of course, the illuminated area grows quadratically with range as it is expanding in two dimensions once divergence of the laser's beam commences. So, the signal return power drops as $1/Range^3$ or $1/Range^4$ depending on whether the target is before or after the commencement of beam divergence. It should be apparent that achieving long range requires emitting a lot of photons.

However, there are limits to the amount of laser power that can be used. Intense near-IR light (800 to 1400nm) can damage vision. As humans cannot see light in this range, we do not blink or avert our eyes to bright near-IR light. But our eyes can focus this light onto our retina. This can result in retinal damage. Longer wavelengths of light, 1400 to 3000nm (or short wave-IR) for example, are absorbed by the aqueous area behind the cornea[3]. So while it is similarly invisible to humans, we can tolerate a lot more laser exposure at those wavelengths – roughly five orders or magnitude more[4]. The reason this is important to understand with regards to direct detection LIDAR is that many LIDAR systems (particularly low-cost automotive LIDAR) use 905 or 940nm as their operating wavelength due to the wide availability of low cost InGaAs-based lasers and Silicon photodiodes. Lasers and photodiodes at short wave-IR tend to be much more expensive, negating the main advantage of direct detection – its simplicity and low cost.

There are other means to improve direct detection range by improving receiver sensitivity. Larger area receiving lenses can be used. Increasing the photon collection area offers improved receiver sensitivity without any additional electronic noise. Doubling the lens diameter offers 4x the receive sensitivity at the expense of a larger and more complex optical system (recall that a 16x increase in gain only translates to a doubling of range). A larger aperture transmit beam can be used to maintain tight collimation of the laser over a longer distance (see the section on Rayleigh range below), but large diameter beams may not be compatible with many scanning methods (small MEMS mirrors, for example). Avalanche photodiodes (APDs) – photodiodes with intrinsic gain – can be used to increase receive sensitivity. As a practical matter they can offer gains of about 5x to 15x before self-generated noise becomes a problem. Avalanche photodiodes tend to be expensive and fragile. They are also generally very small area devices, which complicates the optical design further. Finally, Geiger Mode Avalanche Photo-Detectors (GMAPDs) or Single Photon Avalanche Detectors (SPADs) are available. They offer extreme sensitivity – as little as a single photon is needed for detection. However, once they have been triggered, they require finite time (~5 to 10ηs), to recover before being able to trigger again[5]. While these can make a highly simplified long range LIDAR system, their principle of operation is such that they are susceptible to interference (Solar and adjacent LIDAR systems) and work poorly in snowy, dusty, or foggy environments (a photon reflected off of a snowflake will blind the GMAPD to anything 1.5 to 3m behind the snowflake). As we will discuss later, some applications are not subject to interference from the Sun, adjacent LIDAR systems, or concerned about poor weather environments. In those applications GMAPD based direct detection systems work very well.

Regarding interference, it is also important to note that direct detection systems used in applications where there are other LIDAR systems around (like automotive or autonomous ground vehicles) must design-in some means of interference mitigation. To the receiver of a direct detection LIDAR system, every light pulse at a similar wavelength looks just like its own pulse. This is not a LIDAR exclusive problem. In the early days of automotive radar, pulsed systems were used. Once many cars were equipped with radar, mutual interference became a problem. In response, the automotive radar industry moved to coherent detection techniques – mostly FMCW – largely solving the mutual interference issues[6]. In general, some kind of pulse coding must be used to distinguish "your" laser pulses from other systems. The cost to this is either reduced range (if average laser power is limited due to thermal or eye-safety issues) or a reduced number of spots/second the LIDAR unit is capable of measuring. Pulse coding is difficult to do when using GMAPDs as the time between pulses must be long enough to ensure the GMAPD has recovered from the last pulse.

Finally, it should be noted that direct detection LIDAR does not measure velocity (which can be a valuable input to downstream perception) directly[7]. Velocity may be inferred by measuring target movement across multiple frames; however, this tends to be a low accuracy measurement technique as it depends on repeatable measurements of the target position in each frame. For example, if a target is moving at 15m/s (about 33mph) and frame rate is 20Hz the target would have moved 75cm in one frame. If measurement accuracy is ±10cm (about the best one would expect from a direct detection automotive LIDAR system), then velocity measurement error could be as high as $\pm {}^{10cm}/_{75cm}$ = ±13%. Of course, this could be improved by measuring multiple consecutive frames. But this would take time as measurement accuracy only improves with the square root of the number of measurements taken (for example, 9 averaged measurements improve the accuracy by a factor of 3 while increasing latency by a factor of 9, up to 450ms at the frame rate of 20Hz).

Coherent Detection

Coherent detection involves mixing the incoming light with a sample of the transmitted light (often referred to as the local oscillator). This optical mixing offers two main advantages:

(a) Noiseless amplification through photonic gain achieved via constructive interference. The receive signal is multiplied by the local oscillator. As a result, coherent detection systems achieve excellent sensitivity with very low power lasers.

(b) Mixing the transmit and receive signals effectively results in making the LIDAR system hyper selective. Light that is not at exactly the same wavelength is simply rejected. Sunlight, the major source of noise in direct detection LIDAR, is ignored, as are adjacent LIDAR systems.

While there are several coherent detection schemes, LIDAR frequently uses Frequency Modulated Continuous Wave (FMCW) modulation. So we will use FMCW to illustrate how coherent LIDAR works. The pros and cons of other coherent modulation schemes are mostly similar to FMCW, but where they differ, we will attempt to point out where significant differences lie.

Figure 2 shows a highly simplified graphic description of an FMCW LIDAR system. In this example, the laser operates at around 1550nm and the laser is modulated by a few hundred MHz (from 1550.002 to 1550nm, for example). The transmitted signal (and reflected signal) is centered ~200THz. After optical mixing of the received signal with a sample of the transmitted signal the photodiode is presented with the sum and difference of the two signals (recall that when two signals are mixed together the output is the sum and difference of the two signals: $\cos(a) \bullet \cos(b) = ½ [\cos(a + b) + \cos(a - b)]$). The photodiode is bandwidth limited and unresponsive to the ~400THz sum and detects only the ~few hundred MHz difference signal.

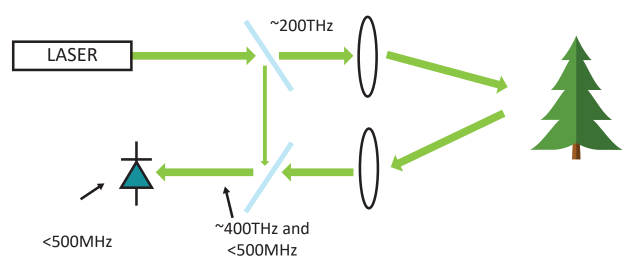
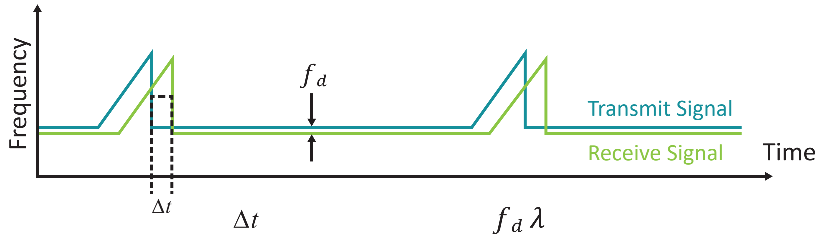


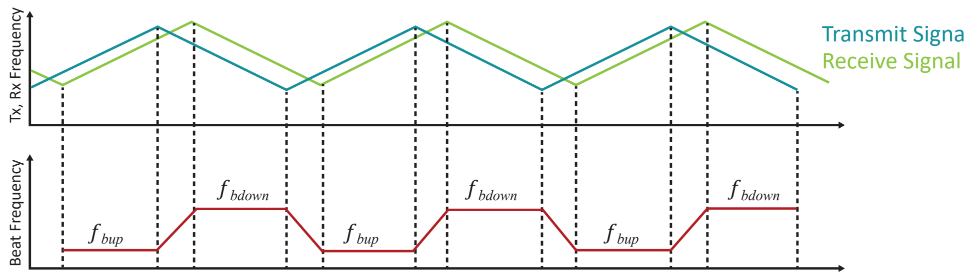
Figure 2. Graphic description of a coherent detection LIDAR system

Range and velocity may be derived by the formulas presented in Figure 3. In practice, the laser is usually swept up and down in frequency, producing triangle waves with range and velocity equations shown in Figure 4.



$$\frac{\Delta t}{2} c = r \qquad \frac{f_d \lambda}{2} = v$$

Figure 3. FMCW range (r) and velocity (v) signals



$$r = \frac{c \cdot T_{sweep}}{4BW_{sweep}} \times \frac{f_{bup} + f_{bdown}}{2} \quad \text{and} \quad v = \frac{c}{2f_0} \times \frac{f_{bup} - f_{bdown}}{2}$$

Where: f_0 $= center\ of\ transmit\ frequency$
T_{sweep} $= sweep\ time$
bW_{sweep} $= sweep\ bandwidth$

Figure 4. Triangle Wave FMCW modulation

In actuality, an FMCW LIDAR system has a lot more intricacies than the simplified system shown in Figure 2. A more complete system is illustrated in Figure 5. Even this is rather simplified from an electronics point of view (signal processing electronics are omitted, for example), but the optical system is fairly complete.

Clearly, this is much more complicated than a direct detection system, so why would LIDAR makers go to so much trouble? Just as we examined in the case of direct detection, the power return equation[8] for coherent LIDAR offers an explanation.

- $Return \propto E_{signal} \times E_{LO} \times \exp[-jt(\omega_{signal} - \omega_{LO})]$ Eq. 2

In Eq. 2 we can see that the return signal is multiplied by the sample taken off the transmitted source (the local oscillator), presented here as E_{LO}. As the path loss for LIDAR is high (per Equation 1), even a small sample of the local oscillator (a few percent) will be much greater than the return signal. The amount of signal amplification is very high, but only for signals at the exact same wavelength. This is where coherent LIDAR gets its high selectivity (interference rejection) from.

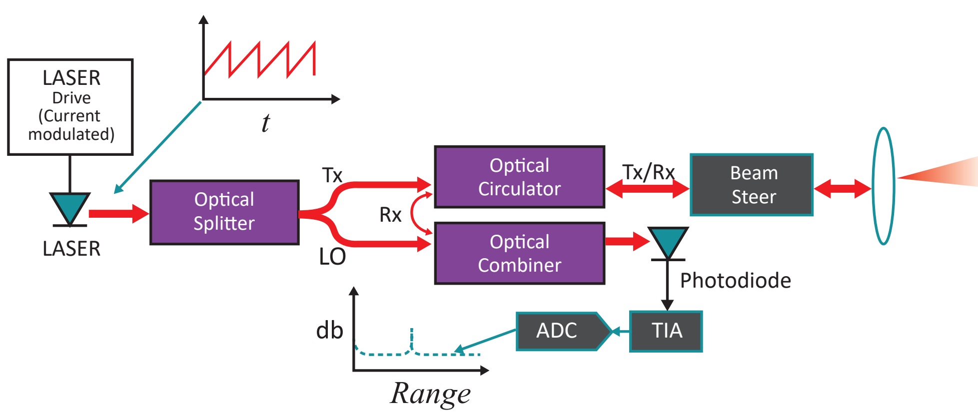


Figure 5. The optical system of an FMCW LIDAR system

As an example of the photon efficiency of coherent LIDAR, an FMCW LIDAR system with ~300m range can be realized with a <200mW laser. Comparable direct detection systems would require 1000x greater peak power for similar range. Many examples of long range coherent LIDAR systems exist serving multiple industries. Some examples include optical altimetry instruments with ranges up to several km and laser Doppler LIDAR instruments for wind characterization with range of >500m (wind is characterized by measuring the velocity and direction of particles in the air – normally just a source of backscatter for most applications). These systems have been in production for some time, so clearly they have achieved Technical Readiness Levels (TRL) of 9.

Another characteristic of coherent LIDAR is that the bandwidth of the signal chain is fairly low. In the previous example (where the laser is swept from 1550.002 to 1550nm) the photodiode bandwidth can be restricted to a few hundred MHz . A direct detection system will normally have as wide bandwidth as possible – often over 2GHz – in order to resolve the leading edge of the receive pulse (even at 2GHz, position resolution is ~15cm). The narrower bandwidth of the FMCW system lowers receiver noise. But the effective bandwidth of an FMCW LIDAR system receive chain is, in use, lower still. Narrower bandwidth allows the designer to use lower noise trans-impedance amplifiers at the photodiodes and slower analog-to-digital converters.

As range increases, receiver acquisition time must be increased to account for the additional round-trip return time. Since a Fourier transform is an integration operation, the receive noise is also integrated for a longer period and is effectively reduced as a result. When coupled with a large aperture transmitter – one capable of maintaining beam collimation over at least much of the system range – we see that the Signal to Noise Ratio (SNR) remains roughly constant as long as the target remains within the Rayleigh range (see below for a brief discussion of Rayleigh range). Figure 6 shows the modeled and actual response of an FMCW LIDAR system operating at 1550nm with a 200mW laser and a 12mm aperture (resulting in ~70m Rayleigh range). We can clearly see that receiver background noise drops with time and that the system SNR is essentially flat within the Rayleigh range. After which the SNR drops at about -40dB/decade - similar to that of a direct detection system. This is advantageous as the need to re-engineer a LIDAR system in response to a somewhat longer range requirement is eliminated. Recall that a 50% increase in LIDAR system range for a direct detection system would require ~3.5 to 5x more laser power (depending on whether the target is within the Rayleigh range or not). In comparison, with a coherent system an increase in beam aperture of ~25% (going from 12mm to 15mm diameter) would suffice.

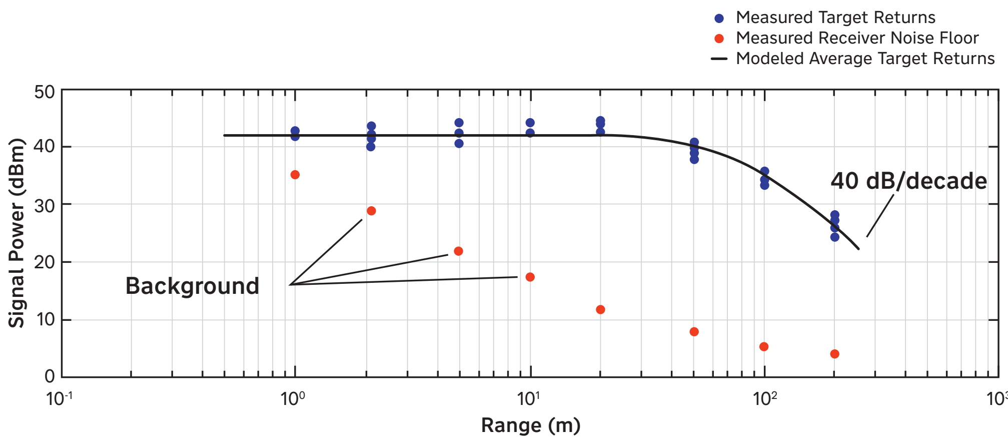


Figure 6. SNR of an FMCW LIDAR system with a 200mW laser operating at 1550 nm and 12mm aperture (based on author's experimental data)

Supplementary Information

Rayleigh Range

We have all noticed that the collimated light from a laser pointer remains a nice, tight, spot that does not vary in diameter over several meters of range. However, at some longer distance the spot begins to grow. What you are seeing is Rayleigh range in action. Rayleigh range is a property of beam divergence. The range at which a collimated beam starts to grow is proportional to its beam waist and wavelength. For a Gaussian beam it is described as $Z_{Rayleigh} = \frac{\pi \omega_o^2}{\lambda}$ where ω_o is the beam radius. At 1550nm a beam radius of 5mm will remain tight out to ~50m after which it begins to expand. By 200m it would be ~20mm in radius. Assuming the same number of photons are still available (no path loss), the optical intensity at 200m would be 16x lower in this example. A beam with 10mm radius would remain tight to ~230m.

Larger beam radius is almost always advantageous for a LIDAR system, however not all beam steering methods can cope with a larger beam. Even those that can often have to be made physically larger to accommodate the larger beam, which might make the overall LIDAR system larger.

But coherent LIDAR is not without challenges. In order to work a laser with long coherence length (or its inverse, narrow line width) is needed[9]. The laser must be able to maintain its phase integrity for long enough for its light to go to, and return from, the farthest target. If the laser's phase changes by more than π radians during transit time coherence may be lost and range measurement ambiguity could result. To make matters worse, this very stable laser must be frequency (in the case of FMCW) or phase (in the case of Phase Modulated Coherent Wave) modulated. Most diode lasers are not up to the task, but recently a number of semiconductor tunable lasers have appeared on the commercial market. Likewise, not every scan mechanism is compatible with coherent detection. There is a need for the receiver to continuously look at each spot for long enough to allow light to go to, and return from, the furthest possible target (recall that we will want to mix some of the transmit signal with the return signal). For a range of 300m, for example, this require ~2µs. For this example, the scan mechanism must remain effectively still for at least 2µs . Many continuously moving scan mechanisms are incapable of this. Finally, it must be noted that the signal processing tasks of coherent LIDAR are significantly greater than direct detection. Fortunately, semiconductor makers have responded with highly capable system-on-chip offerings that integrate data converters, microcontrollers, and DSPs with FFT accelerators to meet these signal processing needs . The Indie Semiconductor iND83301 Surya LIDAR SoC is one such example.

Obscurants

Dust, rain, snow, fog, or other particulate matter are challenging for LIDAR systems. Radar fares much better in obscurant-rich environments as the wavelengths used for radar are much longer (4mm for 77GHz automotive radar, for example) than LIDAR. When signal wavelengths are larger than the size of the obscurants, they tend to bend around them. But LIDAR wavelengths tend to be smaller than water droplets, snowflakes, or dust particles. So some photons simply reflect back as backscatter in foggy, snowy, or dusty conditions. This is a phenomenon well known to anyone who has driven a car on a foggy night. All LIDAR systems struggle in these environments, but direct detection systems that rely on GMAPDs fare particularly badly as they are constantly being driven into avalanche (saturation) by the backscatter. Coherent systems, being much more photon efficient than direct detection, tend to fare better in poor visibility conditions than direct detection. In all cases, it's not that no photons get through the fog (or dust, or snow). Just fewer of them. Coherent LIDARs SNR advantage helps in these conditions.

Blooming

Blooming is common phenomenon for image sensors. We have all seen digital pictures of a scene where bright objects extend a halo around themselves (like a streetlight in a night scene) that washes out some area around it. This is also often seen in LIDAR systems that use image sensors (a GMAPD array, for example) as their receiver. In this case, crosstalk between the pixels often causes blooming when a high intensity retroreflector is in the scene. But blooming can also occur in single pixel receiver systems if the laser aperture is small. In this case the divergence of the laser may illuminate a retroreflector that is a few pixels to the side of the direction the system is looking at. While the method of photon detection (direct or coherent) does not, strictly speaking, influence blooming, direct detection system tend to suffer more from blooming than coherent systems as they tend to have much lower dynamic range than coherent systems. Plus, direct detection systems more frequently rely on image-sensor based receivers (like GMAPD arrays) that are just natively sensitive to blooming.

LIDAR Applications

There are many different LIDAR applications, and their unique requirements often play heavily in determining what photon detection technique is preferred. In this section we will present a few examples to illustrate this as a means of instruction. It should be stressed that this list is not exhaustive and is presented to illustrate where each photon detection technique fits in well.

Aerial Ground Survey

Ground surveying via drones or aircraft mostly relies on photogrammetry, where multiple high resolution photographs are taken from different locations to build a 3D map of the terrain. However, in cases where there is thick forest canopy photogrammetry fails as it cannot see the ground. LIDAR systems using direct detection and GMAPDs work much better in this case as some photons will pierce the thick canopy and hit the ground (and then return). Unlike automotive or autonomous ground vehicle applications in this application there is little risk of solar interference (you are looking straight down) or risk of encountering other LIDAR systems that may interfere with you. Likewise, aerial surveys happen only during favorable weather. Direct detection with GMAPDs is well suited for this application as none of its inherent weaknesses – most notably susceptibility to interference and difficulties in obscurant-laden environments - are problematic here. Figure 7 shows a point cloud of a moderately thickly forested area. Plot a (top) was taken using a GMAPD-based LIDAR system. Plot b (bottom) is the point cloud generated via high-resolution photogrammetry. It's clear in these point clouds that the LIDAR system is capable of sending and receiving photons through the forest canopy and illuminating the tree trunks and ground below.

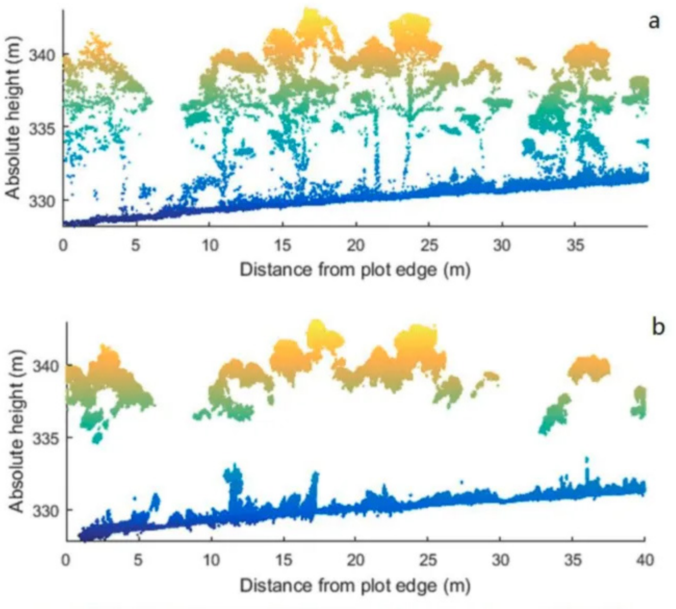


Figure 7. Point cloud of a forest (a) with GMAPD-based LIDAR and (b) with photogrammetry*

Agriculture

Drone mounted LIDAR is used in the evaluation of crop progress. These drones rarely fly higher than ~100m above ground. Solar and adjacent LIDAR system interference is unlikely as the view is straight down. The scene is typically static and there is no velocity information of interest. Lastly, crop monitoring is not performed in bad weather. In such an undemanding application, direct detection LIDAR is adequate for this application.

Automotive/Ground Vehicle Navigation

Ground vehicles (including passenger cars and trucks) generally require ranges of a few hundred meters or less (depending on the maximum speed of the vehicle). However, compared to many other LIDAR applications they face challenging conditions and need to achieve a sufficiently low cost in order to achieve mass-market adoption. They are expected to work without disturbance in environments where many other LIDAR systems are present and there is frequent Solar glare. The environment is dynamic, so loss of a frame due to interference or Solar glare cannot be tolerated. Likewise, they require high probabilities of detection for real targets and very low probability of false alarm. They are power limited (for overall vehicle efficiency and due to tight thermal budget on the LIDAR unit) and must operate reliably in obscurant-rich environments like rain, fog, or snow. These are difficult conditions for direct detection systems – particularly in obscurant-rich environments where signal-to-noise ratio (SNR), and therefore range, is often degraded. Here the chief advantages of coherent detection (photon efficiency and interference immunity) are highly valuable. As shown in Eq. 2, the combined effects of photonic multiplication (for high photon efficiency) and the high selectivity inherent in coherent detection offer an exact match to the needs of ground vehicles (interference immunity from adjacent LIDAR system, and high SNR to overcome the high path loss in foggy/snowy/dusty conditions). The direct, high accuracy, per pixel measurement of velocity inherent with coherent detection helps with and simplifies perception processing-based target separation and identification (for example, a pedestrian versus a cyclist can be easily differentiated using velocity information) while relaxing distance measurement accuracy requirements. The measurement range requirement of a few hundred meters or less also means that extremely long coherence lengths lasers are not necessary, so laser costs are not prohibitive.

Terrain Mapping and Surveying

Mapping and surveying applications often require >2000m range and have fairly slow frame rates as they observe mostly static scenes. They are rarely in danger of being interfered with by other LIDAR systems and even if they temporarily suffer interference, they can simply rescan the scene. The long range needs and unlikelihood of interference makes them good candidates for high power direct detection LIDAR. As they operate outdoors and view scenes parallel to the ground, they can be persistently blinded by low-angle sunlight if using GMAPDs. So traditional photodiodes or APDs are preferred.

Wind Measurement for Wind Turbines

Understanding what the wind direction will be some minutes ahead of time improves the efficiency and reliability of wind turbines. The maximum range needed is typically between 250 to 500m, and the LIDAR must measure wind velocity and direction. The measurement plane is parallel to the ground ±10° or so, so there is a strong possibility of Solar glare. This application is ideally suited to coherent LIDAR with its native velocity measurement capability as well as its high rejection of Solar noise.

Wind Shear Measurement

Several companies build LIDAR-based wind shear measurement systems for safeguarding civilian and military airports (Lockheed Martin and Mitsubishi Electric, for example). The velocity of particles in the air is measured to determine wind velocity over a mapped area. These LIDAR systems typically have ranges of tens of km. As wind shear measurement systems are highly unlikely to encounter other LIDAR systems in their vicinity interference mitigation is not a great concern. As it is costly to make a laser with enough power and sufficient coherence length for such long range, and the natural strengths of coherent detection (photon efficiency and interference rejection) are not needed in this application, direct detection is a natural choice.

Conclusion

Different LIDAR applications benefit from different design approaches. In applications where extremely long range

** Image source: https://wingtra.com/drone-photogrammetry-vs-lidar/*

is needed and there is little risk of LIDAR system mutual interference, high power pulsed direct detection is very suitable. However, for applications like automotive or autonomous ground vehicles where <1km range is required and there is a high likelihood that other, potentially interfering, LIDAR systems are around, coherent detection has several advantages. Most significantly its immunity to interference (including solar), high SNR (important in bad weather conditions), native high accuracy velocity detection to provide additional information to perception systems, and ease of system modification. Coherent signal processing techniques are well known through developments in radar and telecom over the past three decades, and these techniques are equally compatible with LIDAR. The main challenges for coherent LIDAR – the need for tunable, long coherence length lasers and its more mathematically intense signal processing needs (FFTs versus threshold comparators) - have been addressed with the wider availability of appropriate lasers and system-on-chip offerings that integrate most of the electronic signal processing chain.

As an aid to the reader, Table 1 offers a brief summary of performance characteristics for direct detection and coherent LIDAR. While each specific application will require a thorough examination of which performance characteristics are most important, Table 1 offers an indicative quick reference. LIDAR system designers should carefully consider the appropriate photon detection approach for their particular application.

Photon Detection Method	**Range**			**Environment**			**Features**	
	<20m	20 to 1000m	>1000m	Other LiDAR Systems Present	Solar Glare	Obscurants	High Accuracy Velocity Measurement	Low Power
Coherent	+	++		+	++	++	++	+
Direct	+		+		-	--	--	-
Application	**Requirement**							
Aerial Ground Survey			✓					
Agriculture		✓						✓
Automotive & Ground Vehicle Navigation	✓	✓		✓	✓	✓	✓	✓
Terrain Mapping & Surveying			✓		✓			
Wind Shear Measurement			✓			✓		
Wind Heading Measurement		✓			✓	✓	✓	

Coherent Detection Preferred   Direct Detection Preferred

Table 1. Comparative Table, Direct Detection LIDAR Versus Coherent Detection LIDAR

Supplementary Information

Other Aspects of LiDAR System Design: Beam Steering

While we covered coherent versus direct detection in this paper, there are several other design/component choices and tradeoffs when it comes to LIDAR that are not directly tied to the detection method but will greatly influence the architecture of the system. Let's consider one example; the choice of scanning system. The scanning system determines how the transmitted and received beams are directed and steered between each measurement in order to create a 3D representation of the environment – without a way to steer or scan the light beams, the LiDAR would only range a few points in space. Different scanning systems have their own advantages, limitations, and design considerations.

For example, while solid state beam steering solutions are generally desirable due to the resulting smaller system size, perceived reliability and the ability to define arbitrary scanning patterns, these scanners typically are restricted to small apertures, limited field-of-view, and may be costly. On the other hand, mechanical scanners such as polygon mirrors can represent a simpler, cost-effective solution that can scan a wide field-of-view, at the cost of being bulkier and being more prone to wear over time. Furthermore, most beam steering components move in one axis only and creating a full 3D image thus commonly requires the combination of two distinct scanning mechanisms into a hybrid scanning system.

Examples of how the choice of scanning system may influence the design of a LIDAR system

Scan Angle and Field of View

The scanning system's range of achievable scan angles and field of view (FOV) directly impacts the LiDAR's ability to capture data from the environment. Many automotive applications require wide field of views, sometimes as wide as 160 degrees to 180 degrees.

Supplementary Information (continued)

Examples of how the choice of scanning system may influence the design of a LIDAR system	
Scan Patterns	Different LiDAR applications may require specific scan patterns (e.g. single-point scanning, raster scanning, or customized patterns). The choice of scanning system determines the feasibility and ease of implementing these patterns.
Scan Speed and Refresh Rate	Some scanning systems offer high-speed scanning capabilities, enabling faster data acquisition and real-time 3D mapping. At the same time, we saw earlier that it may be desirable to select a scanning system which can dwell at a fixed position before moving, as quickly as possible, to the next position.
Vibration and Shock Tolerance	Some scanning systems are more resilient to vibrations and shocks, which is crucial in applications like automotive LiDAR where shock and vibration is always present.
Pointing Accuracy and Precision	The scanning system's mechanical design, control mechanisms, and feedback systems impact the LIDAR's pointing accuracy and in turn the accuracy and precision of object detection. Factors such as manufacturing tolerances, mechanical stability under vibration and shock as well as position feedback mechanisms play a role in achieving accurate results.
Power Consumption	The choice of scanning system affects power consumption in two ways: First, the power consumption of the scanning mechanism itself, where some mechanical scanners like galvanometric mirrors can require several watts of power, and second, in the losses incurred by the light going through the scanner, which may increase the required transmit power emitted at the source to achieve the required SNR.
Reliability and Maintenance	Reliability and robustness requirements can greatly influence the choice of scanner, as this specification can vary greatly between various solutions, with some potentially being more prone to wear and tear. For example, passenger cars typically have a lifetime of 5,000 to 10,000 hours of operation while some commercial vehicles can require an operational lifetime of up to 50,000 hours.
Integration and Size Constraints	The scanning system's size and integration requirements impact the overall form factor and packaging of the LIDAR unit. Compact scanning systems like MEMS mirrors, solid-state beam-steering or compact polygon mirrors are advantageous in space-constrained applications like Automotive.
Cost	The complexity, components, and control systems associated with different scanning technologies influence the overall cost of the LiDAR system. Furthermore, requirements for fine opto-mechanical alignment or any end-of-line calibration can also highly influence the overall cost of the LIDAR solution.
Environmental Conditions	The choice of scanning system may affect how well the LIDAR system performs in different environmental conditions, such as extreme temperatures and temperature variations, humidity, as well as exposure to dust and dirt.

Ultimately, the choice of scanning system should align with the specific requirements and constraints of the LIDAR application. Designers need to carefully balance factors such as scanning speed, accuracy, size, power consumption, and cost to create an effective and reliable LIDAR solution.

References

1. R. N. Lloyd Espenschied, “A terrain Clearance Indicator,” The Bell System Technical Journal, vol. 18, no. 1, pp. 222 - 234, 1939.
2. G. P. Agrawal, Optical Communication: Its History and Recent Progress, Springer Cham, 2016.
3. Sliney and Wolbarsht, “Safety With Lasers and Other Optical Sources,” Plenum Press, 1980.
4. I. E. C. (IEC), “IEC 60825-2014 Safety of Laser Products - Part 1: Equipment Calssification and Requirements,” IEC, 2014.
5. Brian F. Aull, Erik K. Duerr, “Large-Format Geiger-Mode Avalanche Photodiode,” IEEE JOURNAL OF SELECTED TOPICS IN QUANTUM ELECTRONICS, vol. 24, no. 2, p. 10, 2018.
6. M. Kunert, “The EU project MOSARIM: A general overview of project objectives and conducted work,” in European Radar Conference (EURAD), Amsterdam, 2012.
7. R. N. Mahalati, “ Principles of Sensing for Autonomy,” Stanford University, 2023.
8. S. Bhaskaran, “Direct Detection Time of Flight Lidar Sensor System Design and A Vortex Tracking Algorithm for a Doppler Lidar,” Arizona State University, 2018.
9. M. Harris, G Pearson, “The Role of Laser Coherence Length in Continuous Wave Coherent Laser Radar,” Journal of Modern Optics, vol. 45, no. 8, pp. 1567 - 1581, 1998.

About Pier-Olivier Hamel

Mr. Hamel is Director of Marketing at indie <leading efforts for our LiDAR> program. Prior to joining indie in 2021, he held various roles at LeddarTech including Product Line Manager, product leader and applications engineer. In 2012, he was a product design engineer for embedded software development at Multitel Inc.

Mr. Hamel holds a bachelor’s degree in electrical and electronics engineering from the Université Laval, and completed the Product Management program from the University of California, Berkeley.

About Harvey Weinberg

Mr. Weinberg is currently Director, Sensor Technologies at Microtech Ventures. He previously worked at Analog Devices, a broad line semiconductor company, for 25 years. During the most recent 10 years, Harvey served as Division Technologist for the Automotive Business Unit, principally working on long-time horizon technology identification and early technology development as it pertains to automotive applications. Harvey’s prior roles at ADI have been System Application Engineering Manager for the Automotive Business Unit and, before that, leader of the Applications Engineering Group for MEMS inertial sensors.

Harvey has developed 14 US patents in technologies varying from ultrasonic airflow measurement, to inertial sensor applications, to LIDAR systems. Before ADI, Harvey worked for 12 years as a circuit and systems designer specializing in process control instrumentation.

Harvey holds a Bachelor of Electrical Engineering degree from Concordia University in Montreal, Canada.