A Tutorial and Review of Automobile Direct ToF LiDAR SoCs: Evolution of Next-Generation LiDARs

Kentaro YOSHIOKA†, Nonmember

SUMMARY LiDAR is a distance sensor that plays a key role in the realization of advanced driver assistance systems (ADAS). In this paper, we present a tutorial and review of automotive direct time of flight (dToF) LiDAR from the aspect of circuit systems. We discuss the breakthrough in ADAS LiDARs through comparison with the first-generation LiDAR systems, which were conventionally high-cost and had an immature performance. We define current high-performance and low-cost LiDARs as next-generation LiDAR systems, which have significantly improved the cost and performance by integrating the photodetector, the readout circuit, and the signal processing unit into a single SoC.

This paper targets reader who is new to ADAS LiDARs and will cover the basic principles of LiDAR, also comparing with range methods other than dToF. In addition, we discuss the development of this area through the latest research examples such as the 2-chip approach, 2D SPAD array, and 3D integrated LiDARs.

key words: LiDAR, dToF, ADAS, automotive, SPAD, TDC, ADC.

1. Introduction

Humans have better sensors and decision-making mechanisms than most modern hardware and software systems. Even though we make incredibly childish mistakes from time to time; such mistakes have irreversible consequences, especially when driving a car. In Japan alone, the number of fatalities and injuries in traffic accidents in 2021 reached 2,636 and 361,768, respectively [1]. Since it is impossible to reduce human errors to zero, Advanced Driver-Assistance Systems (ADAS) technologies are developed to compensate for such errors. As a milestone, SAE has set five levels for automated driving, as shown in Table 1 [2]. For example, as of 2022, Level 0-1 automated driving, such as automatic braking and lane-keeping, is available in many commercial vehicles. To name a few examples, Tesla has released a Level 2 partially automated driving function, and Honda also equipped their products with Level 3 equivalent automated driving functions (although the use of such functions are still limited) [3]. Thus, driving functions are being transferred from human drivers to machine-controlled systems in a long-term view.

There are two main approaches to automated driving: 1) driving by constantly recognizing surrounding objects and situations like a human [4], and 2) by utilizing a prerecorded 3D map of the environment and estimating the ego position by fitting the sensed 3D data to the prerecorded map [5]. While the former can drive in an environment without any mapping information, there is a risk that an unexpected misdetection, e.g. overlooking the gore points, may cause a severe consequence. However, while the latter requires recorded mapping information, it can mitigate the above accidents by labeling the possible driving region in advance. In particular, a high-quality depth sensor is essential for the latter automated driving system. In addition, the high resolution distance sensor can be used to recognize surrounding cars and pedestrians by detecting point cloud objects [6, 7], simultaneously obtaining distances between objects.

We summarize the typical depth sensors used in ADAS systems in Table 2. It is important to note that there is a trade-off between performance and cost for all sensors. Unlike millimeter-wave (mmwave) radar, stereo cameras, and ultrasound sensors, Light Detection And Ranging (LiDAR) sensors require a mechanical component for scanning, which makes LiDAR more expensive than other sensors. On the

<table>
<thead>
<tr>
<th>Level</th>
<th>Controller</th>
<th>Backup controller</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Human</td>
<td>Human</td>
<td>Automatic brakes</td>
</tr>
<tr>
<td>1</td>
<td>Human/AV</td>
<td>Human</td>
<td>Lane centering</td>
</tr>
<tr>
<td>2</td>
<td>AV</td>
<td>Human</td>
<td>Adaptive cruise control</td>
</tr>
<tr>
<td>3</td>
<td>AV</td>
<td>Human</td>
<td>Automated driving under limited conditions</td>
</tr>
<tr>
<td>4</td>
<td>AV</td>
<td>AV</td>
<td>Driverless taxi</td>
</tr>
<tr>
<td>5</td>
<td>AV</td>
<td>AV</td>
<td>Can drive anywhere</td>
</tr>
</tbody>
</table>

AV: Automated vehicle
other hand, despite their cost, LiDARs have attracted attention because they are the only depth sensors that can provide high-resolution and long-range measurements.

Next, the requirements of the distance sensors for automated driving are discussed. Since 150m is the braking distance of a car travelling at 120km/h on a freeway, a distance sensor capable of sensing 200m is required for forward monitoring. This target is very challenging, regarding that the VLP-32 [8] has a maximum distance of 50m and does not meet the requirements of the use of the highway. Furthermore, when driving in urban areas, it is essential not to overlook pedestrians at a distance. Thus, depth detection with high resolution is required (for example, a horizontal resolution of 0.1–0.2 degrees), which is very difficult to achieve with ultrasonic or radar sensors [9, 10].

It is also crucial to have the capability to sense in all weather conditions (e.g., extreme sunlight, rain, snow, fog) to increase the reliability of automated driving. Among Table 2, the mmwave radar is a sensor that is not easily affected by the weather. On the other hand, when a LiDAR is placed in a foggy environment, its effective range is shortened because the laser is scattered. It is challenging to build a reliable automated driving system with a single sensor for these reasons. Thus, it will be necessary to take a sensor fusion approach where the sensors compensate for each other’s weaknesses [11, 12].

This paper reviews the development of automotive LiDAR for automatic driving from the aspect of circuit systems. Although a number of conference presentations have been made in recent years, to the best of our knowledge, there has been no comprehensive review paper on LiDARs for automatic driving. The target readership of this paper is intended to be at the introductory level of ADAS LiDAR, and we set the goal to obtain a general overview of the research direction of the field. The main focus will be on dToF LiDAR with scanning mechanisms using 850–950nm lasers, which are expected to be mass-produced for automotive applications as of 2022 [13–26]. Therefore, 1550nm LiDAR [27], FMCW LiDAR [28, 29], Flash LiDAR [30–32], and iToF LiDAR [33–36] are mentioned as comparisons, but are not the main targets of the review. In addition, the discussion will focus as much as possible on the LiDAR circuit system, especially the photodetector (PD), readout circuitry, and signal processing, and the discussion of optics and lasers will be kept to a minimum.

We organize this paper as follows: in Chapter 2, we explain the measurement principle of the dToF LiDAR and the issues unique to ADAS LiDARs. Then, in Chapter 3, the first-generation LiDAR is studied in detail. Although the first-generation LiDAR made a significant contribution to automated driving prototypes, its high cost faced challenges to mass production. Next, in Chapters 4 and 5, we will discuss how the latest LiDARs, which we call next-generation LiDARs, have made a breakthrough in cost and performance. Finally, Chapter 6 summarizes and discusses the prospects of the field.
be more suited for short-range applications requiring high measurement performance for ADAS, and iToF may.

For these two reasons, it is challenging to achieve the 200m single-photon avalanche diode (SPAD). As a result, the PD an avalanche photodiode (APD) instead of a highly sensitive response to capture the modulated laser, it is necessary to use In addition, since the photodetector must have a linear response but shorter measurement distances. accuracy, since higher sensor modulation frequency leads iToF holds a trade-off between measurement distance and phase shift and allows higher precision [33–36]. However, to precise measurement but shorter measurement distances. For example, the number of laser photons returned is 1/16 for 200m distance measurement compared to 50m, making the operation very difficult. Moreover, sunlight is the most significant noise source for outdoor LiDARs: for automotive applications, the LiDAR must function under extreme sunlight. Fig 2 illustrates such harsh operating conditions, where the sunlight-triggered outputs can be larger than the laser in long-range measurement conditions.

In principle, LiDAR ranging can be expressed in terms of SNR, where the signal is defined as the number of returning laser photons and noise as the number of noise photons (mainly sunlight) input in a certain unit of time [16].

\[
\text{LiDAR SNR} = \log_{20} \frac{\text{Number of laser photons}}{\text{Number of noise photons}} \quad (2)
\]

Another restriction for automotive LiDAR is that the emitting laser power must comply with eye safety requirements. For automotive applications, it is common to comply with the strictest class-1 eye safety, i.e., the laser must not harm the human eye under any circumstances. In other words, due to the strict laser power limitations, the signal power cannot be further increased. On the other hand, eq.(2) shows that optical filters that filter sunlight and an increased sensitivity of PDs can contribute to SNR.

### 2.2 Basic LiDAR architectures

The operating principle of Fig.1 is known as direct ToF (dToF), which is the main method for automotive LiDAR. On the other hand, the indirect ToF (iToF) method modulates the laser and measures the distance by detecting the phase shift and allows higher precision [33–36]. However, iToF holds a trade-off between measurement distance and accuracy, since higher sensor modulation frequency leads to precise measurement but shorter measurement distances. In addition, since the photodetector must have a linear response to capture the modulated laser, it is necessary to use an avalanche photodiode (APD) instead of a highly sensitive single-photon avalanche diode (SPAD). As a result, the PD sensitivity is inevitably lower than that of the dToF method. For these two reasons, it is challenging to achieve the 200m measurement performance required for ADAS, and iToF may be more suited for short-range applications requiring high precision, such as robotics [37].

The dToF LiDARs can be categorized into two types: flash [30–32] and scanning. As shown in Fig.3, the flash emits a laser beam over the entire field of view, and a 2D array of PDs receives the reflected light, similar to image sensors. The advantage of the flash method is that the LiDAR is free of mechanical parts, resulting in a low cost and a high frame rate. However, when the number of pixels in the flash LiDAR is \(N \times M\), the laser power \(P\) will be diffused to \(N \times M\) pixels, resulting in a weak laser power of \(P/(N \times M)\) per pixel. Therefore, while high resolution is easy to achieve with flash LiDARs, long-distance measurement such as >20m is difficult. Therefore, the flash LiDAR’s potential applications are short-range LiDARs attached to the side of a car or robotics application. On the other hand, scanning LiDARs obtain \(M\) pixels and perform a horizontal/vertical scan. Thus, the laser power per pixel is \(P/M\), which is much better than the flash. On the other hand, the cost and frame rates degrade due to the scanning procedure.

There are also several types of LiDAR scanning methods: rotating mirror [8,26], polygon mirror [13,16,18], and MEMS mirror [20,21,23]. The rotating mirror is extremely bulky but generally has good optical properties (less laser attenuation) and can obtain 360-degree data. The polygon mirror obtains data by raster scanning within the FoV, as shown in Fig.4. In order to perform such scans, two actuating mirrors are utilized for both the laser and the receiver optical path. Thus, the implementation becomes bulky due to the mounting of the mirror itself and the motor driving the mirror. Finally, the MEMS mirror eliminates the need for bulky mechanical components by using an extremely small movable micromirror. Therefore, the size of the LiDAR can be significantly reduced and is sometimes referred to as solid-state LiDAR. On the other hand, the MEMS mirror has poor optical properties; the trade-off is the degradation in LiDAR SNR.

### 3. First Generation LiDARs

Velodyne’s rotational LiDAR [8, 38], shown in Fig.5, is implemented by stacking the laser and receiver boards vertically. We define such LiDAR as first generation LiDAR in this paper, since it was released in the infancy of ADAS
LiDAR. However, such LiDARs achieved high-quality depth sensing and played a key role in many self-driving prototypes [5].

As shown in the schematic in Fig.6(a), the APD is used as the PD of the first-generation LiDAR. The APD output is amplified by TIA and VGA and then quantized by a high-speed ADC, and the ToF was calculated in the digital processor. The first generation LiDAR can be seen as an integration of "point" measuring distance sensors consisting of a laser and PD pair. On the other hand, such implementation required many discrete components. As a result, the cost of the first generation LiDAR was very high and fragile. In addition, it was difficult to scale the performance on the same body because the increase in resolution directly impacted the number of components. Moreover, although APDs are more sensitive than ordinary photodetectors, they were inadequate for long-distance measurements.

4. Next Generation LiDARs

The first-generation LiDAR advanced the horizon of automated driving and greatly expanded the ADAS market. However, it was challenging to adopt this technology in mass-produced vehicles without lowering costs and extending its performance. To achieve these goals, research on next-generation LiDAR has been carried out. In this paper, we discuss the fundamental concept of the next-generation LiDAR as the integration of SPAD, readout circuit, and the signal processing circuit.

Illustrated in Fig.6(b), ref. [15] is a breakthrough in LiDAR SoC, which achieves the integration of the SPAD array, readout circuit, DSP and memory in a single chip. For automotive LiDAR, the number of SPADs per pixel is several 10s to mitigate the signal saturation (or quenching time) as described below and as a result, the total number of SPADs in the array is in the order of 100-1000. Thus, connecting these SPADs to the signal processing circuit is quite challenging. Therefore, ref. [15] designed a SoC using a high voltage CMOS process to integrate 384 SPADs, readout circuits, DSPs, and memory on a single chip.

In contrast to the first-generation LiDARs that used many discrete components, this SoC achieves the same function on a single chip, paving the way for significantly low-cost LiDARs. In addition, as is known as Moore’s law, CMOS scaling is expected to increase the number of transistors with a lower cost, making the LiDAR performance scalable.

4.1 SPAD detectors

Both APDs and SPADs operate photodiodes with a strong reverse bias, but SPADs are very sensitive photodetectors capable of single-photon detection [39–43]. SPADs are biased above the breakdown voltage (20-30V in silicon) to operate the diodes in Geiger mode. In Geiger mode, when a photon is received, the amplification ratio of the device ideally becomes infinite, allowing the detection of single photons by flowing a large current independent of the photon intensity. On the other hand, if such a current continues to flow, the device will be destroyed. Therefore, negative feedback is applied by the accompanying quenching resistor to stop the current forcibly as in Fig.8.

The SPAD contributes to the long-range performance of the LiDAR because its strong amplification enables the detection of single photon and faint returning lasers during long distance measurements. The key SPAD design parameter is the photon receiving probability and directly relates to device sensitivity (i.e. quantum efficiency or photon detection efficiency (PDE)). The other parameters are the quenching time and the number of SPADs assigned to each pixel. The higher the PDE, weaker lasers can be detected, directly affecting long-distance performance. In addition, the quenching time is closely related to sunlight resistance. If the quenching time is long, the SPAD cannot respond to the desired laser photon if the SPAD is recharging from the ignition by noise photon such as sunlight and degrade LiDAR performance. The best way to reduce the quenching time is to lower the quenching resistance, but this is typically a trade-off against device reliability. In general, multiple SPADs are utilized in each pixel to mitigate the pile-up by providing redundancy. Even if one SPAD fires, another SPAD can receive the laser signal.

4.2 TDC based readout circuitry

While the output of an APD is an analog quantity proportional to the intensity of light, the SPAD output can be treated as a digital pulse by shaping the output with a buffer (as in Fig.8 $V_{out}$). Moreover, dToF LiDARs can sufficiently calculate the ToF from the time between the laser emission and the pulse rise. Furthermore, while it is difficult to implement several tens or hundreds of high speed ADCs on an SoC due to its small area, a time-to-digital converter (TDC) circuit is used in [15], which is a circuit specialized to measure time. TDC was initially introduced as a time quantization
Fig. 6  (a) System diagram of first-generation LiDAR. (b) System diagram of next-generation LiDARs.

Fig. 7  Fully integrated LiDAR SoC [15] ©IEEE

Fig. 8  Schematic and the operation diagram of the SPAD with passive quenching. When a photon enters the SPAD, an avalanche current flows in the SPAD. At the same time, a potential difference in the quenching resistor causes the SPAD bias voltage to drop, the Geiger mode automatically stops. While larger quenching resistor provides better reliability, the recharge time prolongs and results in a longer dead-time.

Fig. 9  TDC readout mechanism in [15] ©IEEE

circuit for digital PLL and returns the digital value of the time difference between the two inputs [44,45]. Compared to ADCs, TDCs can be composed almost out of digital circuits, and by distributing the reference clock signal to a large number of TDCs, an array of TDCs can be realized with a small area. In addition, the available time resolution of TDCs is as high as 10-100 ps, achieving ToF accuracy that cannot be easily achieved with ADCs.

4.3 Signal processing circuits

The SoC integration made room for a richer signal processing, leading to the development of signal processing techniques specific to dToF LiDAR. One of the most popular signal processing methods is signal accumulation. As shown
in Fig. 9, by accumulating the results of the $N$ TDC measurement in the same situation and obtaining a histogram, accumulation can improve the SNR by $\sqrt{N}$. While sunlight is a random event with no correlation, laser light is a deterministic event and is observed at the same time. By using accumulation, the SNR improves as the number of measurements is increased, but it is a trade-off with FPS since it takes more measurements to obtain a single pixel.

In addition, in [15], a certain threshold for TDC activation is used to increase sunlight tolerance. The problems caused by sunlight are: 1) the histogram memory become gigantic if all incoming sunlight events are recorded, 2) due to the finite TDC reset time, the TDC may miss the laser if the sunlight triggers the TDC. Therefore, by adding a threshold to the TDC trigger (e.g., triggering TDC only when four SPADs fire simultaneously), we can solve both problems simultaneously.

Finally, Fig. 10 shows the point cloud image obtained by the prototype LiDAR reported in [15], where the integration of SPAD, readout circuit and signal processing circuit realized a high-performance LiDAR capable of recognizing walls up to 100m away.

5. Next generation LiDAR SoCs

In the previous chapter, we studied the evolution of the next generation LiDAR based on [15]. This chapter will discuss more advanced research examples in detail to deepen our understanding of the recent field of ADAS LiDARs.

5.1 2-chip approach

Firstly, we will introduce the 2-chip approach. The most distinctive feature of [15]'s architecture is that it integrates both digital and SPADs on the same chip, which is an excellent choice in terms of cost. However, it poses a challenge when further extending the performance. For example, a special diode structure is required to obtain the best performance from SPADs, which is difficult to achieve in advanced digital CMOS processes. Therefore, [15] utilizes a legacy 180nm CMOS for the SoC. If we can adopt a more advanced CMOS process, the signal processing capability and the time resolution of the TDC can be dramatically improved by Moore’s law.

Ref. [16]’s LiDAR adopts the 2-chip approach, which achieves the best of the two worlds by adopting the most suitable process technology for the SPAD and DSP (300nm and 28nm), respectively. However, the number of output wires for SPADs is significant, and simply separating the chips will cause wiring problems. To solve this problem, [16] uses a SiPM configuration, which connects multiple SPADs (60 in [16]) in parallel and extracts the output as a summed current to mitigate wiring complexity. Additionally, when using a TDC, a comparator with a set threshold is used to convert the output to pulses, allowing the same processing as the conventional TDC-based systems.

In [15], the LiDAR SNR was improved by histogramming the TDC output of multiple measurement results. However, since the TDC is not triggered unless the SPAD’s firing exceeded a certain threshold, the accumulation was not effective when the returning laser was very weak at a long distance. If we can directly accumulate the raw SPAD waveform, the LiDAR can effectively utilize the information below the accumulation threshold, but such setup will require...
an ADC as the readout circuit.

The LiDAR SoC [16] aims to further improve LiDAR performance by adopting a hybrid readout circuitry, which switches between ADC and TDC for far and short distances, respectively (Fig. 11). Fig. 12 illustrates the concept of hybrid readout; a TDC is utilized for short-distance measurements, since high distance resolution is required and the returning laser SNR is sufficiently high. In contrast, at long distances (>20m) where SNR is severe, the SiPM waveform is read directly by an ADC and the accumulation is performed at the raw waveform level. Notably, the distance resolution requirement is relaxed at longer distances so that even an ADC of 400 MS/s can achieve sufficient distance measurement performance. As a result, the TDC/ADC hybrid architecture significantly reduces the ADC speed requirement and minimizes the hardware cost. [16] achieves 200m distance measurement for the first time, owing to the ADC-based waveform accumulation and custom SPAD integration.

5.2 2D SPAD array approach

Both [15] and [16] required a scanning mechanism to perform raster scanning (Fig. 4) and used SPADs arranged in a 1D array. Although raster scanning reduces the required SPADs, it involves scanning the receiving (RX) and transmitting (TX) laser beams. Since the RX scanning mechanisms are much larger than those of the TX due to the larger aperture ratio, this posed a significant challenge when down-scaling the LiDAR size. [19] proposed an in-sensor scanning method in which RX raster scanning is performed within the 2D SPAD array (Fig. 13), thus eliminating the need for RX scanning machinery. The removal of the bulky RX scanning system significantly reduces the LiDAR sizing (Fig. 14). In addition, [19] uses an active quenching technique that resets the SPADs with transistors to shorten the quenching time. This enabled the reduction of SPADs per pixel and [19] achieved a higher resolution LiDAR system at less cost.

5.3 3D integration approach

Previous studies extended their performance by fabricating SPAD and DSP separately with suitable processes, respectively. On the other hand, as the number of pixels increases, the interchip connectivity becomes complex and such approaches lose scalability. Moreover, due to the use of SiPMs, it was necessary to use an ADC with low area efficiency to achieve 200m range performance [16, 18].

To address this issue, [23, 24] proposed a LiDAR SoC with the 3D integration approach (Fig. 15, 16). The 3D integration allows SPAD and DSP chips to be fabricated in their suitable process (90nm and 40nm, respectively). Fur-
thermore, the large number of high-density 3D interchip connections allows the wiring of 100,000 SPADs; since a simple SPAD instead of a SiPM can be used, area-efficient TDC can be adopted as a readout circuit. In addition, another breakthrough is the use of microlens and backside illumination (BSI) technology on SPADs, which is a commonly used technique to increase the sensitivity of image sensors. By applying the microlens and BSI to the SPAD, [23, 24] showed that the PDE can be dramatically improved to 22%, at a wavelength of 905 nm, which is almost a 2x increase from conventional SPADs. Such an increase in PDE lead to significant improvement in LiDAR performance.

Finally, we compare the performance of the LiDARs mentioned above in Table3. It is difficult to directly compare the performance of these works, since they all have different scanning methods, optics, laser power, and resolutions. For example, the SNR of a MEMS mirror is much lower than that of a rotating mirror, which is a disadvantage for LiDAR long-range performance (however, they accomplish much smaller LiDAR sizing). Therefore, rather than the absolute value of the LiDAR performance, it is necessary to evaluate the advancement of the circuit and system technology in this field.

6. Conclusions and future prospects

A tutorial and review of LiDAR for dToF ADAS were presented, where LiDARs are key distance sensors upon realizing automated driving systems. First, we discussed the breakthrough in next-generation LiDARs through comparison with the first-generation LiDAR systems. Next-generation LiDAR systems significantly improved their cost and performance by integrating the photodetector, the readout circuit, and the signal processing unit into a single SoC. In addition, we discussed the latest developments in this field by discussing the newest research examples such as the two-chip approach, 2D SPAD array, and 3D integration LiDAR.

There are two main directions for the future development of LiDAR: commercial and research. DToF automotive LiDAR will extend its performance by evolving the SPAD and DSP performance through 3D integration and extensive signal processing. When the mass producibility of DToF LiDAR reaches a sufficient level, such LiDARs will be installed in ADAS systems of commercial vehicles.

As for research, the 1550nm LiDAR still has excellent potential. For example, dToF LiDARs have the risk of being spoofed by malicious attackers [47, 48], and there are high hopes for FMCW LiDARs that can prevent this from happening [49]. In addition, 1550nm LiDARs with silicon photonics can realize solid-state laser scanning to further scale the cost, and research development attracts significant attention [29].

Acknowledgements

This work was supported in part by the JST CREST program under Grant JPMJCR21D2, and in part by the Japan Society for the Promotion of Science (JSPS) KAKENHI, under Grant 21K20413.

References

[16] Kentaro Yoshioka, Hiroshi Kubota, Tomonori Fukushima, Satoshi
<table>
<thead>
<tr>
<th>Table 3 Performance comparison of first-generation and next-generation LiDARs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SPADs</strong></td>
</tr>
<tr>
<td><strong>Optical System</strong></td>
</tr>
<tr>
<td><strong>Pixel-Resolution</strong></td>
</tr>
<tr>
<td><strong>Laser Wavelength [nm]</strong></td>
</tr>
<tr>
<td><strong>Laser average power [mW]</strong></td>
</tr>
<tr>
<td><strong>Laser repetition rate [MHz]</strong></td>
</tr>
<tr>
<td><strong>FPS</strong></td>
</tr>
<tr>
<td><strong>Target Reflectivity [%]</strong></td>
</tr>
<tr>
<td><strong>Background light [klux]</strong></td>
</tr>
<tr>
<td><strong>Distance range [m]</strong></td>
</tr>
<tr>
<td><strong>1 sigma error @max distance</strong></td>
</tr>
<tr>
<td><strong>SoC power consumption [W]</strong></td>
</tr>
</tbody>
</table>