You Li and Javier Ibanez-Guzman

Lidar for Autonomous Driving

The principles, challenges, and trends for automotive lidar and perception systems

©ISTOCKPHOTO.COM/OONAL

Autonomous vehicles rely on their perception systems to acquire information about their immediate surroundings. It is necessary to detect the presence of other vehicles, pedestrians, and other relevant entities. Safety concerns and the need for accurate estimations have led to the introduction of lidar systems to complement camera- or radar-based perception systems. This article presents a review of state-of-the-art automotive lidar technologies and the perception algorithms used with those technologies. Lidar systems are introduced first by analyzing such a system's main components, from laser transmitter to beamscanning mechanism. The advantages/disadvantages and the current status of various solutions are introduced and compared. Then, the specific perception pipeline for lidar data processing is detailed from an autonomous vehicle perspective. The model-driven approaches and emerging deep learning (DL) solutions are reviewed. Finally, we provide an overview of the limitations, challenges, and trends for automotive lidars and perception systems.

Introduction

Autonomous driving is entering a preindustrialization phase, with significant progress attained over the past years. Sensors initially capture data representations of the environment, which are processed by perception algorithms to build the vehicle's immediate environment used for autonomous vehicle navigation. A perception system for autonomous vehicle navigation consists of a combination of active and passive sensors, namely, cameras, radars, and lidars [1]. Lidars are active sensors that illuminate the surroundings by emitting lasers. Ranges are measured precisely by processing the received laser returns from the reflecting surfaces. As stated in [2], lidars are "poised to significantly alter the balance in commercial, military, and intelligence operations, as radar has done over the past seven decades."

Despite much progress in camera-based perception, image processing methods estimate distances. This approach encounters difficulties when estimating distances for cross-traffic entities, particularly for monocular solutions. The 2007 DARPA

Digital Object Identifier 10.1109/MSP.2020.2973615
Date of current version: 26 June 2020

Grand Challenge, a milestone in autonomous driving, demonstrated the potential of lidar perception systems. The top three teams were all equipped with multiple lidars. A 64-layer lidar, the Velodyne HDL64 [55], played a critical central role for the winning and runner-up teams [3], [4].

Currently, many high-level autonomous vehicles use lidars as part of their perception systems despite their high cost and moving parts. A typical example is their incorporation into the autonomous vehicles being tested as robovehicles in different countries, e.g., the Eco-Mobility by Autonomous Vehicles in the Paris-Saclay area (EVAPS) field operational test in France [56], where several participants (e.g., Renault, Transdev, Vedecom, and so forth) collaborate to exploit mobility services based on autonomous vehicles in the Saclay area of Paris. One of the prototypes developed by Renault is shown in Figure 1. As a result, there are more than 20 companies developing distinctive lidar systems for autonomous driving systems, ranging from low level to high end, a sort of "big bang." Which lidar type(s) will dominate autonomous driving in the future remains to be seen.

On the other hand, lidar-based algorithms also entered into a fast track. For an autonomous vehicle, lidars are mainly used for perception and localization. Due to space constraints, this article focuses only on perception usages. In the context of autonomous driving, a perception system provides a machine-interpretable representation of the environment around the vehicle. From a user's perspective, the output of a perception system comprises the following three levels of information.

- *Physical description*: the pose, velocity, and shape of objects
- *Semantic description*: the categories of objects
- *Intention prediction*: the likelihood of an object's behavior.

Therefore, the lidar outputs are used for the object detection, classification, tracking, and intention prediction, corresponding to the various layers of information. Due to lidar's superiority in ranging accuracy, the provided physical information is highly reliable. Although the semantic information carried by lidar is more or less difficult than that which is acquired from a camera, a contextual sensor is good at object recognition. In practice, lidars are combined with cameras to complement each other [5]: a camera is poor in distance estimation, while lidar is inadequate for object recognition. Precise physical and semantic information, together with map information, will improve intention prediction without any doubts. With many years of progress, a lidar-centric perception system will mature for model-based processing algorithms; however, emerging DL methods are changing this domain. Traditional model-based lidar data processing methods are computation friendly and explicable. Data-driven DL methods have demonstrated extraordinary capabilities in providing semantic information, which is the weak point of traditional methods.

Lidar technologies

A typical lidar operates by scanning its field of view (FoV) with one or several laser beams. This is done through a delicately designed beamsteering system. The laser beam is generated by an amplitude-modulated laser diode that emits at near-infrared (NIR) wavelength. The laser beam is reflected by the environment back to the scanner, with the returned signal received by a photodetector. Fast electronics filter the signal and measure the difference between the transmitted and received signals, which are proportional to the distance. The range is estimated from the sensor model based on this difference. The difference in variations of reflected energy due to surface materials as well as state of the milieu between the transmitter and receiver are

FIGURE 1. The autonomous vehicle prototype developed by Groupe RENAULT for the EVAPS project. The most evident sensor is the Velodyne UltraPuck lidar on top.

compensated for through signal processing. The lidar outputs include 3D point clouds that correspond to the scanned environments and the intensities that correspond to the reflected laser energies. Figure 2 shows a conceptual representation of this operating principle.

A lidar system can be partitioned into the laser rangefinder system and the scanning system. The laser rangefinder comprises 1) the laser transmitter, which illuminates the target via a modulated wave; 2) the photodetector, which generates the electronic signal from the reflected photons after optical processing and photoelectric conversion; 3) the optics, which collimate the emitted laser and focus the reflected signal onto the photodetector; and 4) signal processing electronics, which estimate the distance between the laser source and the reflecting surface, based on the received signal. The scanning system typically will steer laser beams at different azimuths and vertical angles, denoted by ϕ_i, θ_i, where i is an index that determines the direction at which the beam is being pointed.

This section initially addresses the principles of a rangefinder to understand its measurement process and limitations, then, it introduces the scanning systems that define the sensor FoV. It is then possible to classify lidars based on the technologies they use. This classification is then applied to examine the commercially available automotive type lidars.

Laser rangefinder principles

A rangefinder that measures the distance to an object by a laser beam is known as a *laser rangefinder*. The manner in which they operate depends on the type of signal modulation used in the laser beam. Pulsed lasers are used so that their time of flight (ToF) can be measured, these are known as *direct-detection* laser rangefinders. The laser signal can also be a frequency-modulated continuous wave (FMCW), which indirectly measures the distance and velocity from the Doppler effect. These are known as *coherent-detection* laser rangefinders.

Lidar power equation

The transmitted laser is first attenuated through the transmission medium, then diffused as it reflects from the target surface. It is partially captured by the receiving optics and finally transformed into an electrical signal by a photodetector. For a target at distance r, the amount of received power P_r by the photodetector based from a pulsed-laser emitter can be approximately modeled as [6]

$$P_r = E_p \frac{c\eta A_r}{2r^2} \cdot \beta \cdot T_r, \tag{1}$$

where E_p is the total energy of a transmitted pulse laser and c is the light speed. A_r represents the area of receive aperture at range r, η is the overall system efficiency, and β is the reflectance of the target's surface, which is decided by the surface properties and the incident angle. In a simple case of Lambertian reflection with a reflectivity of $0 < \Gamma < 1$, it is given by $\beta = \Gamma/\pi$. The final part, T_r, denotes the transmission loss through the transmission medium. When the lidar works under adverse conditions (e.g., fog, rain, dust, snow, and so on), the particles in the air scatter and absorb the photons. Equation (1) reveals that the received power p_r decreases quadratically with respect to distance r: an object at hundreds of meters away is order of magnitudes "darker" than at tens of meters. Simply increasing the power of the laser transmitter is restricted by eye-safety standard IEC 60825 [7]. To overcome this, the overall system efficiency must be improved through optics, photodetectors, and more advanced signal processing algorithms. For an FMCW laser, (1) still holds, except for slight differences.

ToF

A ToF laser rangefinder measures the range by calculating the time difference between the transmitted and received lasers

$$r = \frac{1}{2n} c\Delta t, \tag{2}$$

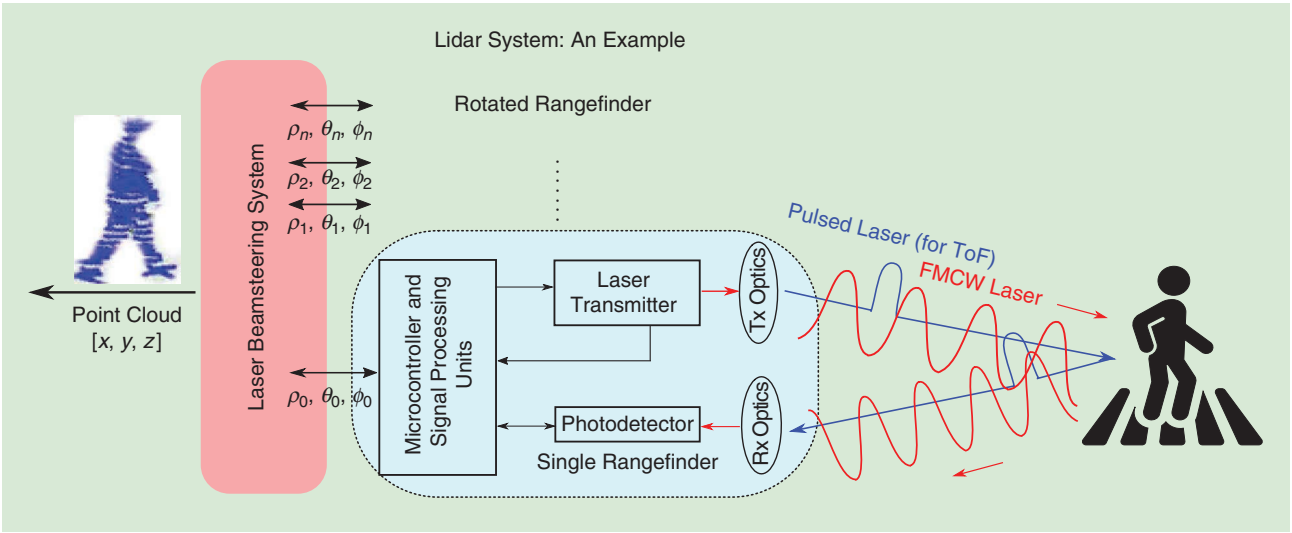

FIGURE 2. An example of a ToF laser rangefinder. The rangefinder uses either a direct or coherent method to measure the distance at a certain direction controlled by the scanning system. Tx: transmitter; Rx: receiver.

where c is the light speed and n is the index of refraction of the propagation medium (≈ 1 for air). Δt is the time gap between the transmitted and received lasers. ToF lidars prevail in the current automotive lidar market because of their simple structure and signal processing methods. However, the potential for increasing their maximum range is constrained by limited transmit power due to eye-safety requirements. In ToF lidar, the return signal can be interference from strong sunlight or laser beams from other ToF lidars.

Coherent detection

By mixing the local carrier signal with the received signal, it is possible to demodulate the received signal and thus it is possible to obtain the phase and frequency shift of the laser signal, hence acquiring the distance and velocity from the reflecting surface. This can be regarded as the optical version of FMCW radars, which are popular in today's passenger vehicles' advanced driver assistant systems. FMCW lidars continuously emit a frequency-modulated laser signal (e.g., linearly chirped laser) to a target, while keeping a reference signal, also known as a *local oscillator*. The common modulation functions are sawtooth or triangle waves.

Because a FMCW lidar continuously illuminates objects using less emitted power for this purpose, thus complying with eye-safety requirements and opening the possibility to use more power to extend their FoV. The signals used for coherent detection are shown in Figure 3. The intermediate frequency [(IF) in red] can be generated by mixing the local oscillator signal (a linearly chirped triangular modulation function) from the laser transmitter (in green) with the laser signal reflected from observed surfaces (in blue). The processing of the resulting signal generates IF f_{if}, shown in red in Figure 3. By assuming that the Doppler frequency shift f_d is less than IF f_{if}, we have

$$f_{\text{if}} = \frac{4rB}{ct} = \frac{f_{\text{if}}^{+} + f_{\text{if}}^{-}}{2}, \quad f_d = \frac{f_{\text{if}}^{+} - f_{\text{if}}^{-}}{2} (\text{when} \quad f_d < f_{\text{if}}), \quad (3)$$

where B, r, and t are the modulation bandwidth, waveform period, and light speed, respectively. The velocity is obtained as

$$v = \frac{f_d \lambda}{2}, \quad (4)$$

where λ is the laser wavelength.

FMCW lidar is able to directly measure the distance and velocity at the same time, while for ToF lidar, the speed is obtained from indirect estimation through several consecutive sensor readings. By using a FMCW laser signal, it is possible to reduce the interference effect from other laser sources and strong sunlight; however, FMCW lidars require high-quality laser generators that possess long coherent distances.

Laser transmission and reception

The generation of the laser signals and their emission as well as the receiver electronics of the reflected signals also characterize the performance and cost of the laser rangefinders.

Laser sources

ToF lidars need a pulsed- (amplitude-modulated) laser signal, which is generated using a pulsed-laser diode or a fiber laser. A diode laser causes laser oscillation by flowing an electric current to the diode's junction. Diode lasers can be grouped into two classes: edge-emitting lasers (EELs) and vertical-cavity surface-emitting lasers (VCSELs). EELs have been applied in the telecommunication industry for a long time. VCSELs output a circular beam, while EELs transmit an elliptical laser beam, requiring additional beamshaping optics. In a VCSEL, forming a 2D laser array on a single chip is easier than for an EEL, which is important because it increases the lidar resolution. By contrast, the range for a VCSEL is shorter due to power limitations.

The pulsed-laser diodes used in automotive applications are hybrid devices; that is, a laser chip is mounted with capacitors that are triggered by a MOSFET transistor. Thus, at every gate opening, the electric charge accumulated in the capacitors will be discharged into the chip, which emits the optical pulse in a controlled manner. These sources are cost-effective, as their 905-nm output can be detected by economical silicon (Si) detectors; however, these diodes have a limited pulse-repetition rate and lower peak power, possibly requiring cooling. Laser diode sources for 3D flash lidar use diode stack technology, with several edge-emitting bars assembled into a vertical stack. Heat dissipation becomes an issue, hence the need for heat sinks as well as the accumulation of emitted power beyond eye-safe requirements. Fiber lasers can have higher output power, which is useful when operating at high wavelengths. Their output beams can be split and routed to multiple sensor locations using optical fiber, resulting in better pulse-repetition frequency, better beam quality, and so forth. However, they can be bulky and thus resulting in noncompact systems that are difficult to be integrated in vehicles.

Laser wavelength

Selecting an appropriate wavelength of laser should have a comprehensive consideration of atmospheric windows, eye-safety requirements, and cost. The 850–950-nm (NIR) and

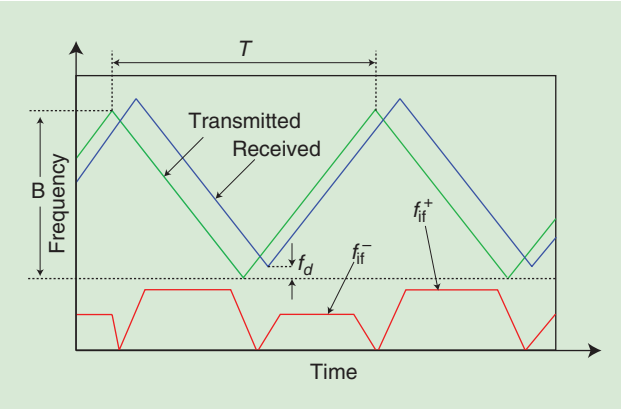

FIGURE 3. The principle of coherent detection: the distance is estimated by the intermediate frequency (red line) generated by mixing transmitted and received light waves [8].

1,550-nm [short-wave infrared (SWIR)] lasers are mostly utilized because of their popularity in industry. Either a low-price diode laser or a more powerful fiber laser at a wavelength of 850–950 nm or 1,550 nm is easily purchased from the market. The maximum power permitted by eye-safety standards for a 1,550-nm laser is higher than that of lasers in the 850–950-nm range, which means a larger range could be achieved. However, expensive indium gallium arsenide (InGaAs)-based photodiodes are required to detect laser returns at 1,550 nm. The efficiency of InGaAs-based photodiodes is lower than that of mature Si ones for NIR lasers. In addition, the atmospheric water absorption for 1,550 nm is stronger than that of 850–950 nm; therefore, lidar systems at NIR wavelengths (905 nm, for instance) are still commonplace.

Photo detector

A photo detector converts optical power to electrical power using the photoelectric effect. Photosensitivity that describes a photodetector's response when receiving photons is one of the most critical characteristics. The photosensitivity depends on the wavelength of the received laser; therefore, selecting a photodetector for a lidar system is closely related to the choice of laser wavelength. The most popular detectors are p-i-n photodiodes, avalanche photodiodes (APDs), single-photon avalanche diodes (SPADs), and Si photomultipliers (SiPMs).

P-i-n photodiodes

These diodes are formed by a p-i-n junction, which creates a depletion region free from mobile charge carriers. By applying a reverse bias to a photodiode, absorbing a photon will generate a current flow in the reverse-biased photodiode.

APD

An APD is a photodiode that applies reverse voltage to multiply photocurrent through the avalanche effect. The APD's ability to multiply signals reduces the effect of noise and achieves a higher internal current gain (roughly 100) and signal-to-noise ratio (SNR) than does the p-i-n photodiode. As a result, APDs are quite common in contemporary lidar systems. Si-based APDs are sensitive through the visible spectral region until the NIR is approximately 1,000 nm. At longer wavelengths up to 1,700 nm, InGaAs APDs are available, although at a higher cost.

SPAD

A SPAD is an APD that is designed to operate with a reverse-bias voltage above the breakdown voltage (Geiger mode), which allows the detection of very few photons in a very short time. SPADs can achieve a gain of 10^6, which is significantly higher than that of APDs; this characteristic allows the SPAD to detect extremely weak light at long distance. Furthermore, the CMOS technology used for a SPAD fabrication enables an integrated array of photodiodes on one chip. This is desirable for increasing lidar's resolution, while cutting the cost and power consumption.

SiPM

An SiPM is based on a SPAD and enables photon counting. The Geiger mode, in which a SPAD operates, is a photon-trigger mode, from which a SPAD cannot distinguish the magnitude of received photo flux. To overcome this issue, the SiPM integrates a dense array of "microcells" (a pair of SPADs and a quench resistor) working identically and independently. The SiPM's output is, in essence, a combination of the photocurrents detected from each microcell. Using this approach, an SiPM is capable of giving information on the magnitude of an instantaneous photon flux.

Scanning system

A scanning system (or beamsteering system) is designed to enable the transmitted lasers to rapidly explore a large area. The existing scanning approaches are usually classified as either mechanical spinning or solid state. The former usually contains a bulky rotating mirror system like the Velodyne HDL64 [55] used in the early stages of autonomous driving history. *Solid state* refers to a scanning system without moving parts (even though some are still steered by micromirrors), which is preferred by the automotive industry.

Mechanical spinning

Currently, the most popular scanning solution for automotive lidar is the mechanical spinning system [10], which steers the laser beams through a rotating assembly (e.g., mirror, prism, and so on) controlled by a motor to create a large FoV. Conventionally, nodding-mirror and polygonal-mirror systems [11] are the main types applied. For example, for the mechanical spinning scheme shown in Figure 4(a), an embedded nodding-mirror system tilts the lasers to generate a vertical FoV. Then, a 360° horizontal FoV is achieved by rotating the lidar base. State-of-the-art lidars use multiple beams to reduce the movable mechanism. For instance, the Velodyne VLP series uses arrays of laser diodes and photodiodes to increase point-cloud densities. The mechanical spinning system offers the advantage of a high SNR over a wide FOV; however, the rotating mechanism is bulky for integration inside a vehicle and is fragile in harsh conditions such as during vibration, which is quite common in automotive applications. A typical product example is Velodyne's HDL64.

Microelectromechanical systems microscanning

Microelectromechanical systems (MEMS) technology allows for the fabrication of miniature mechanical and electromechanical devices using Si fabrication techniques. In essence, a MEMS mirror is a mirror embedded on a chip [12]. The MEMS mirror is rotated by balancing between two opposing forces: an electromagnetic force (Lorentz force) produced by the conductive coil around the mirror and an elastic force from a torsion bar, which serves as the axis of rotation. This principle is shown in Figure 4(b). The MEMS mirrors can either be single axis for 1D movement [13] or dual axis for 2D movement. Also, a MEMS mirror can work in resonant mode at its characteristic oscillation frequency to obtain a large deflection angle and high operating frequency. In nonresonant

mode, a MEMS scanning mirror can be controlled to follow a programmed scan trajectory. For example, for a MEMS-based AEye lidar, the lidar can dynamically change the FoV and scanning path to focus on some critical parts. Although MEMS lidars still contain moving parts, this near-solid-state technology is still promising because the mature techniques in the integrated circuit industry are able to satisfy the strict cost requirements.

Flash

Originally applied for spacecraft in autonomous landing and docking with satellites, 3D flash lidars [14] totally remove the rotating parts within scanning systems. Hence, they are truly solid state. A flash lidar behaves as a camera, in that a single laser is spread by an optical diffuser to illuminate the whole scene at once. Then, it uses a 2D array of photodiodes (similar to the CMOS/charge-coupled device for the camera) to capture the laser returns, which are finally processed to form 3D point clouds, as shown in Figure 4(c).

Because all the pixels of flash lidar measure the ranges simultaneously, the issue of movement compensation caused by platform motion is avoided. In addition, the semiconductor-based 3D flash lidars facilitate fabrication and packaging for massive production, which leads to lower cost. However, the critical issue of 3D flash lidar is its limited detecting range (usually <100 m), because a single diffused laser is responsible for detecting the whole area under a small power threshold for eye safety. Another disadvantage is its limited FoV because it cannot rotate and scan the surroundings the way a scanning-type lidar does.

Optical phased array

As a type of true solid-state lidar, optical phased-array (OPA) lidars [15], [16] do not comprise moving components. Similar to phased-array radar, an OPA is able to steer the laser beams through various types of phase modulators. The speed of light can be changed using the optical phase modulators when the lasers are passing through the lens, as illustrated in Figure 4(d). Consequently, different light speeds in different paths allow for

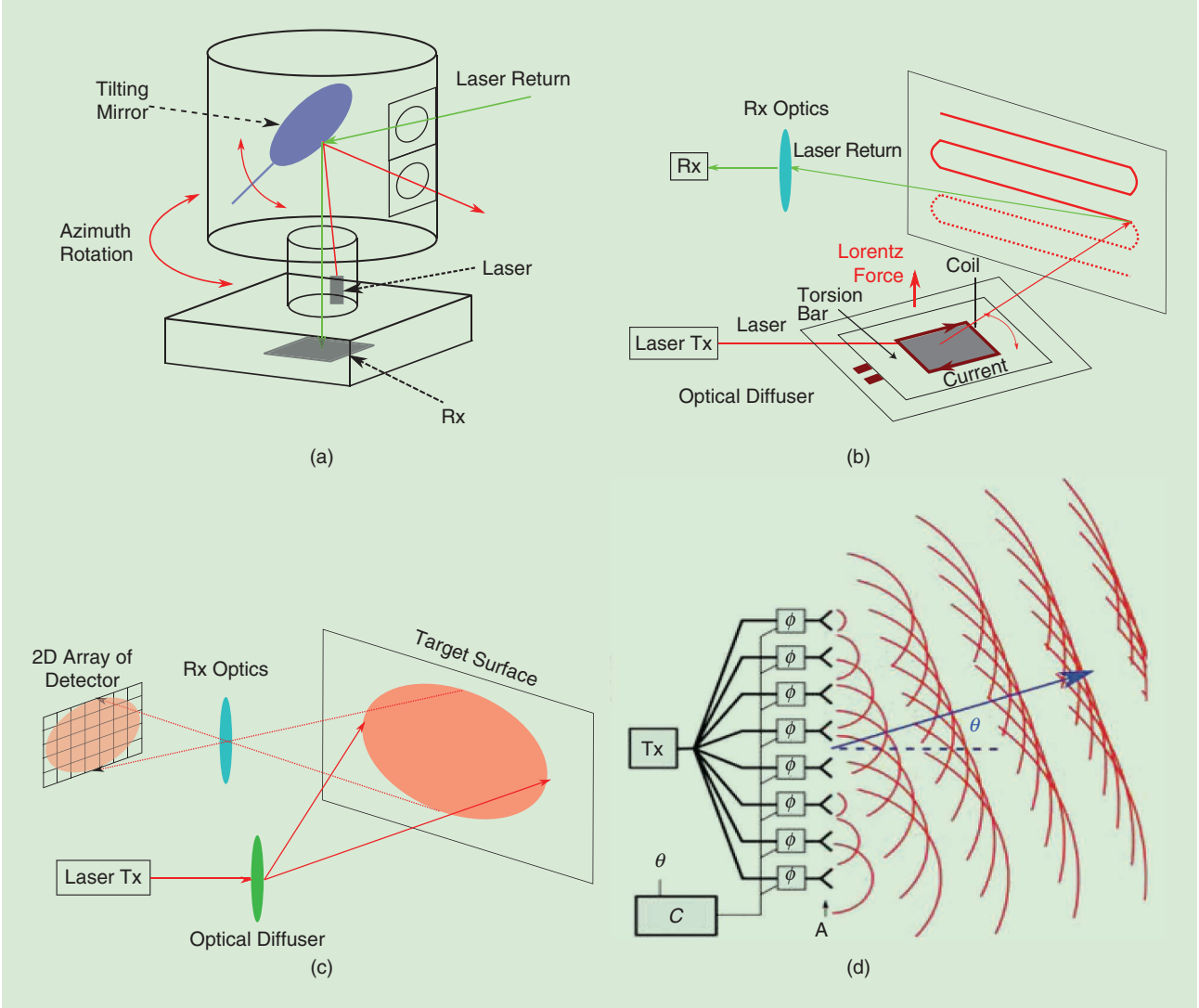

FIGURE 4. Lidar systems categorized by scanning approaches. (a) The principle of mechanical spinning lidar, (b) the principle of microelectromechanical systems lidar, (c) the principle of flash lidar, and (d) the principle of OPA lidar [9].

control of the optical wavefront shape and hence, the steering angles. Although OPA was once seen as a promising technology, there is not yet a commercial product in the market.

Current status of automotive lidar

Mechanical spinning lidar is the first venture into mass-produced cars. Announced in 2017, Audi released its latest luxury sedan, the A8, which equipped Valeo's SCALA lidar for automated driving functions, the first commercially available vehicle carrying automotive-grade lidar in the world. Valeo's SCALA [57] is a four-layer mechanical spinning lidar similar to its cousin, the IBEO Lux4. Empowered by SCALA, the A8 is able to achieve L3-level automated-driving functions without needing hands on the steering wheel (which requires legislative approval). In 2019, Valeo gained a €500 million order of its next-generation lidar, the SCALA2, from several car manufacturers.

At the same time, to reduce cost and improve robustness, many companies focus on solid-state scanning systems. As listed in Table. 1, Innoviz, Continental, and Quanergy are developing MEMS, flash, and OPA lidars, respectively. In 2018, Bavarian Motor Works announced a collaboration with Innoviz for series production beginning in 2021. To increase the maximum detection range, some in the industry have used SPAD arrays working in a single-photon detection mode (Geiger mode). The Ouster OS-1 64 [58] adopted CMOS-based SPADs to detect 850-nm lasers emitted by a 2D VCSEL array. Toyota made a lidar prototype containing a CMOS SPAD array (202 × 96 pixels) for receiving 905-nm lasers [17]. Princeton Lightwave, Inc. (acquired by Argo.ai) also realized a SPAD lidar prototype [59], although little information has been disclosed. As for SiPMs, a lidar prototype [18] was made by SensL (acquired by OmmniVision), while the commercial products are still under development.

Some companies have switched to SWIR lasers (e.g., 1,550 nm) because of its higher power threshold allowed than that of NIR lasers such as Luminar (who announced a collaboration with Toyota) and AEye. Coherent detection-based FMCW lidars are being targeted by car manufacturers and investors as well. Strobe and Blackmore, two representative FMCW lidar start-ups, were quickly acquired by Cruise and Aurora, respectively. In Table 1, we classify and list several representative automotive lidar suppliers and their disclosed technologies.

Another lidar trend is that of overcoming adverse weather conditions, such as rain, fog, snow, dust, and so forth [19], [20]. According to (1), adverse weather conditions increase the transmission loss T_r and weaken the reflectivity of an object β so that the received energy becomes less. Because SWIR lasers (e.g., 1,550 nm) can achieve higher transmission power, lidars belonging to this wavelength are expected to have better performances in harsh weather.

Lidar perception system

For an autonomous vehicle, its perception system classifies the perceived environment into hierarchical object descriptions (i.e., physical, semantic, and intention awareness) from the perception sensor outputs, its localization, and map data. As depicted in Figure 5, a traditional pipeline [4], [21] of processing lidar data consists of four steps: object detection, tracking, recognition, and motion prediction. The recent rise of DL technologies is changing this classic flow, and we will introduce it after the classic approaches. Due to the popularity of Velodyne lidars in research communities, the reviewed data processing methods are mainly based on this mechanical spinning lidar.

Object detection

Object-detection algorithms extract the object candidates and estimate their physical information, namely, the positions and shapes of the detected objects. Because in most traffic scenes the targets are perpendicular to flat ground, object-detection algorithms usually comprise ground filtering and clustering. Ground filtering labels a point cloud either ground or nonground. Then, nonground points are grouped into different objects using clustering methods.

In early research [4], the point clouds from lidar are projected into polar grids, subdividing 360° around the lidar. The points inside each grid cell are treated consecutively to generate a virtual scan, which specifies the region as free, occupied, and occluded. Occupied virtual scans are grouped into object clusters. The authors in [21] followed this method while utilizing a grid-based local plane-fitting approach instead of processing every point, as was done in [4]. The grids that are able to be fitted as a plane are classified as ground grids, and the remaining nonground grids are clustered by the connected component labeling. However, the polar grid-based methods always need a projection of 3D lidar points into discrete grids, which lose raw information from lidar measurements.

Processing lidar signals in spherical coordinates (r, φ, θ) provides a better approach. In this article, we used the Velodyne UltraPuck, where the vertical angle for each laser beam is fixed and the azimuth angle is decided by the scanning time and

Table 1. Representative lidar manufactures and the adopted technologies.

		Mechanical Spinning	MEMS	Flash	OPA	Undisclosed
ToF lidar	NIR	Velodyne, IBEO, Valeo, Ouster,* Hesai, and Robosense	Innoviz Robosense	Continental Xenomatix	Quanergy	
	SWIR	Luminar	AEye and Hesai	Argo* (Princeton Lightwave)		
FMCW lidar					Cruise (Strobe)	Aurora (Blackmore, 1,550 nm)

*The manufacturers that utilize single-photon Geiger-mode SPAD as a photodetector.

motor speed. Therefore, every range reading can be represented by $P_{i,j} = (\rho_{i,j}, \varphi_i, \theta_j,)$, where i refers to a certain laser beam and j is the azimuth-angle index, as shown in Figure 6. This approach naturally fills the range readings into a predefined data buffer (range image) and thus allows for fast access to a point and its neighbors. Processing lidar data in range view has become popular in recent years; for instance, based on a range image, the authors in [22] segmented the ground points in each column. The remaining nonground points are grouped easily through criterions of distance and angle. For a 32-beam lidar, they reached 4 ms in an Intel i5 processor.

Zermas et al. [23] processed the range image row by row. They applied the clustering in each scan line (actually, the row in the range image) and then merged the clusters scan line by

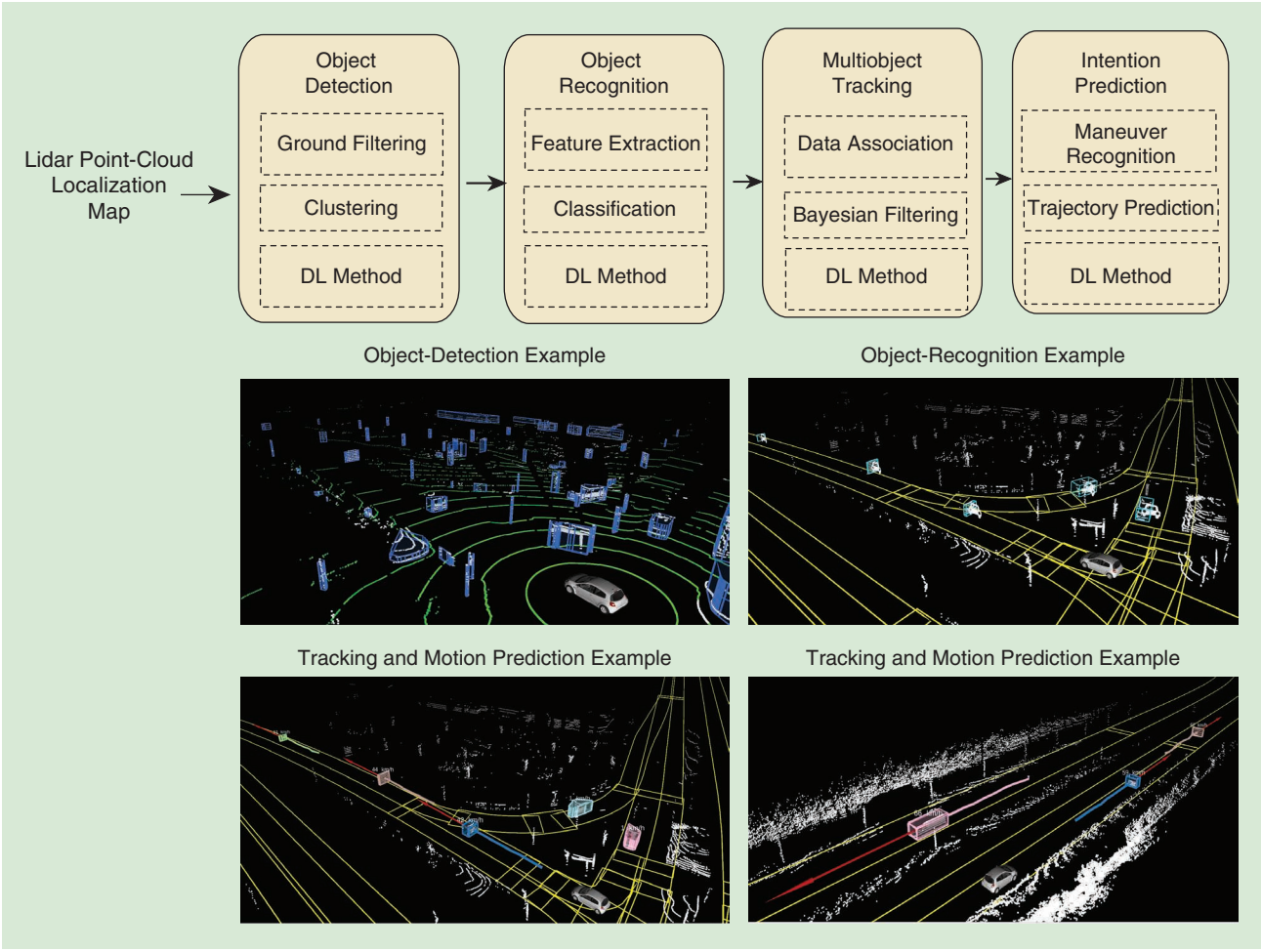

FIGURE 5. The pipeline of a classic lidar perception system, with the exemplary outputs of each step. The examples are from the authors' platform, as shown in Figure 1. Note that after object detection, we only process the objects within the road (as denoted by the yellow lines).

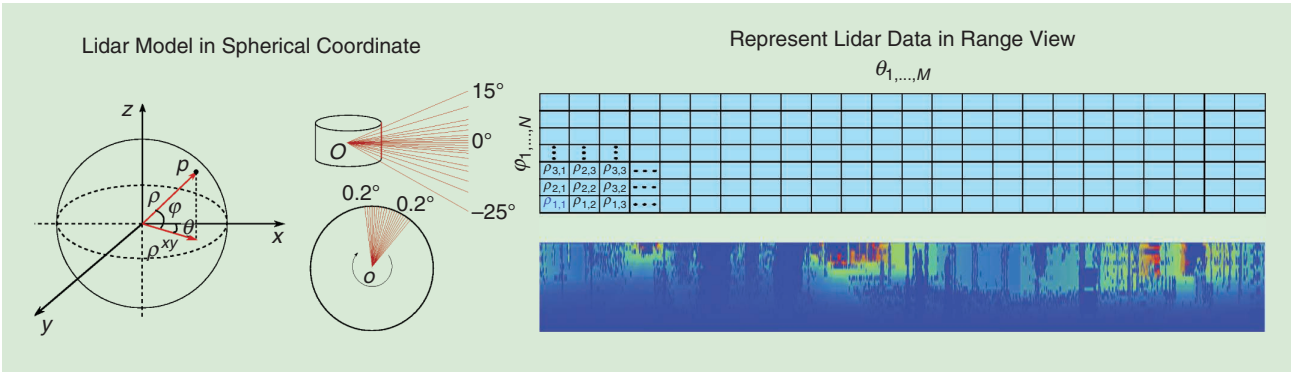

FIGURE 6. The range view of spinning lidar (Velodyne UltraPuck) for further processing. The range image (32 × 1,800) in pseudocolor facilitates the subsequent processing.

scan line. Figure 5 shows a sample result of ground filtering and clustering from our implementation based on the range image as well. The green lines are ground points, and the nonground points are grouped into object candidates (in blue polygons). Object detection provides the initial physical information, e.g., the position of an object. The following steps, such as recognition and tracking, complement semantic and more physical information, e.g., the heading and speed to the detected objects.

Object recognition

Machine learning (ML)-based object-recognition methods furnish the semantic information (e.g., the classes of pedestrian, vehicle, truck, tree, building, and so on) to the detected objects. A typical recognition process employed in [21] comprises a feature-extraction step, calculating compact object descriptors, and a classification step, where pretrained classifiers predict the categories of objects based on the extracted features. As summarized in [24], the features proposed in literature can be roughly divided into two classes: the global features for the whole object or the local features for each point. An object's size, radius, central moments, or maximum intensity [21], [25] are the most basic global features.

Applying principal component analysis (PCA) in 3D point clouds is another effective method used to acquire global shape features. As adopted in [26], three salience features (surfaceness, linearness, and scatterness) can be acquired by analyzing the eigenvalues acquired from the PCA. As for local features, the authors in [21] calculated the three salience features for each point and its neighboring points. Three histograms, each containing four bins spaced between 0 and 1 for the three salience features, are extracted as local features. A more complicated feature is the spin image (SI), introduced by Wang et al. [27]. An SI is created by spinning a grid around the normal surface n of a given point p. The virtual pixels of an SI are the distances to either line through the n or to the plane defined by p and n. The authors in [26] transformed this individual pointwise feature into a global feature: for an object, only the SI of its central point is utilized as an object descriptor. In the literature, there are more sophisticated features, such as a global Fourier histogram [24] descriptor; however, real-time requirements restrain the complexity of its features.

After feature extraction, the following classification is a typical supervised ML process: a classifier trained by a ground-truth data set predicts the class of input objects. Well-known data sets such as KITTI [60] provide abundant resources. From the arsenal of ML, plenty of ML tools such as naive Bayes [25], support vector machines (SVMs) [21], [24], [28], k-nearest neighbor (NN) [26], random forest (RF), and gradient boosting trees [29] can be applied. An SVM with a radial basis function (RBF) kernel is still the most popular method due to its speed and accuracy. Figure 5 shows the recognition results on the detected on-road objects based on our implementations (an SVM with an RBF kernel). Recently, Capellier et al. [30] applied an evidential neural network to classify the lidar object. Evidential classifiers can better handle the unknown classes that are frequently encountered in practice.

Object tracking

Multiple object-tracking (MOT) algorithms correlate and locate the detected/recognized objects through spatiotemporal consistency. MOT maintains the identities of detected objects and yields their physical states, e.g., trajectories, poses, and velocities. MOT is a classic engineering problem [31] that has been researched for a long time. A basic architecture mainly includes a single-object tracker, which "optimally" estimates the state of the tracked object, and data association, which assigns new detections to the trackers.

A single-object tracker models the movement as a dynamic state-space model and estimates its state under the Bayesian filtering framework. The Kalman filter (KF) family, the classic KF under Gaussian-linear assumption and its variants, the extended KF (EKF), and the unscented KF (UKF) make up the popular toolbox. The authors in [32] employed a KF with a constant velocity model to track lidar detections. As a nonlinear version of KF, EKF is utilized for lidar object tracking in [33]. Extending the single dynamic model to multiple maneuver models, the interacting multiple model (IMM) filter is able to handle more complicated cases. The IMM filter consists of several filters running in parallel, and each filter uses a different motion model. For a single object, an IMM-UKF filter was applied in [34], where three UKFs work for three motion models: constant velocity, constant turn rate, and random motion.

As another common approach, a particle filter (PF) is designed for more general cases that do not meet the Gaussian-linear assumption. The application of PFs in lidar data processing can be traced back to the 2017 DARPA Grand Challenge [4], where a Rao–Blackwellized PF was used. However, PFs require a large number of particles, especially for high-dimensional state space. Hence, the KF family is more popular in real-time perception systems.

Data association connects the detections with the tracks. The simplest method used is the NN filter (implemented in [32]), which assigns the detections to their closest tracks based on the Euclidean or Mahalanobis distance between the detection and the track. The NN filter is insufficient for clutter scenarios. In contrast, the joint probabilistic data association filter (JPDAF) offers a soft, probabilistic approach for detection-track association. JPDAF considers all of the possible detections (including no detection) in a gating window and estimates their assignment probabilities to the tracks, taking the weighted average of all the association hypotheses. In [34], JPDAF was applied for data association and an IMM-UKF filter was used to track an individual object.

In contrast to radar-based MOT, in which all of the detections are usually modeled as points, lidar-based MOT is distinctive in that it should track the shapes of detections as well. The simplest shape model is a 2D bounding box [4], which assumes that the detections are car-like objects. L-shape fitting [35] is the most common approach used to estimate the bounding box's center, width, height, and heading; however, a 2D bounding box is insufficient for more general objects, such as a pedestrian, tree, building, and so on. A more sophisticated method [36] implemented multiple shape models: points,

polygons, L-shape, and lines for various objects. When tracking a moving object, its shape varies with the changes of pose and sensor viewpoint. Kraemer et al. [37] implemented a tracking method that simultaneously estimates the states of both poses and shapes represented by 2D polylines.

Object-intention prediction

The previously introduced modules provide the past and current information of detected targets. While in autonomous driving systems the decision-making and path-planning algorithms require future motion of the tracked targets, previous works based on certain kinematic models that are assumed to perfectly fit the detected objects are not applicable for long-term predictions. To address this shortcoming, maneuver or behavior recognition is proposed based on ML methods. Common maneuvers for a vehicle are cutting in, changing lanes, braking, overtaking, and so forth. The authors in [38] modeled the behaviors of car following and lane changing using a Gaussian mixture model (GMM) or hidden Markov model (HMM). Based on the maneuver classification realized by the HMM, the authors in [39] predicted the vehicles' motions by VGMMs (variational GMMs) under the constraints of vehicle-interaction models. With the success of recurrent neural networks (RNNs) in modeling temporal sequential data, long short-term memory (LSTM)-based methods are becoming more popular. Phillips et al. [40] used LSTM to classify the drivers' intentions at intersections, and the results show that LSTM outperforms other traditional ML methods. The authors in [41] proposed an encoder–decoder LSTM model to recognize maneuvers and predict trajectory. Beyond recognizing the maneuvers of a single object, social LSTM [42] was proposed to capture the interactions of all the objects. This is achieved by social pooling, which downsamples a target's neighboring object's LSTM states into a social tensor. The authors in [43] and [44] applied and improved the original social-pooling part for the purpose of vehicle-trajectory prediction.

Emerging DL methods

After huge successes in computer vision and speech recognition, waves of DL arrived in lidar data processing as well. DL [45] is a subset of ML algorithms that mainly uses multilayer neural networks. In contrast to traditional ML methods such as SVMs, DL technologies are able to automatically extract features from the raw input. Convolutional neural networks (CNNs) and RNNs such as LSTM are the most frequently used tools.

The basic components of perception system, ground segmentation, object detection, tracking, and recognition can all be realized using deep neural networks (DNNs). For instance, in [46], ground points were segmented by applying CNNs to lidar points represented by multichannel range images. In contrast to the object detection based on clustering in which arbitrary targets can be detected, DNN-based solutions achieve object detection using recognition, thanks to the paradigm of supervised learning. As in [47], vehicles can be detected using CNN-based neural networks in a bird's-eye view (BEV) representation of lidar points. A more complex neural network was proposed in [48], where CNNs are utilized in both a range image and BEV of lidar data and then fused with camera detections. However, due to the physical limitations of lidar, only vehicles can be effectively detected by lidar, with the best achieving results of only 52.4% average precision pedestrian detection in a KITTI benchmark of (using the DENFIDet method when writing this article) [61].

The authors in [49] integrated evidential theory into DL architecture for lidar-based road segmentation and mapping. Object tracking has been realized by DL as well. Different from the tracking-by-filtering framework described in traditional tracking algorithms, the authors in [50] proposed a deep structural model under tracking by a detection framework. A detection net first processes a sequence of lidar data and images to generate detection proposals. Then, tracks are estimated by finding the best associations of detections, which are achieved by a marching net and scoring net.

Apart from improving the traditional perception components, pointwise semantic segmentation, which was hard to realize before, is now achievable using deep learning. A well-known method, PointNet, was proposed in [52] to semantically segment 3D point clouds for general use. However, due to the sparsity of lidar data with respect to distance, the method does not work well for autonomous driving scenarios. SqueezeSeg [53] achieved real-time segmentation by applying CNNs in the range view of lidar points. Due to a lack of massive annotated data sets, the performances of these two methods are not ready for deployment in real situations, although this situation has been changed by SemanticKITTI [54], [62], the latest and biggest pointwise annotated data set based on KITTI. Based on this data set, RangeNet [51] demonstrated a fascinating performance and speed from a simple DNN structure. Figure 7 shows its structure and a sample result of RangeNet. With more annotated data sets, we have sufficient reason to expect lidar-based semantic segmentation will have better performance.

Conclusions and future directions

In this article, a review of lidar technologies was presented. How a lidar "sees" the world and what constitutes a lidar were introduced, and the main development directions of lidar technologies were analyzed as well. In summary, current automotive lidars face the following constraints or challenges: 1) cost; 2) meeting automotive reliability and safety standards (e.g., ISO 26262 and IEC 61508); 3) a long measuring distance (e.g., >200 m for highway applications); 4) adverse weather conditions, e.g., rain, fog, snow, and so on; 5) image-level resolution; and 6) its smaller size, which facilitates integration. At present, all of the possible solutions, varying from laser sources (905 nm versus 1,550 nm), scanning methods (spinning/MEMS/OPA/flash), or ranging principles (ToF or FMCW) have been exploited to overcome several or all of these difficulties. Although it is very difficult to predict which automotive lidar solution(s) will dominate in the future, one thing is certain: automotive lidars are walking out of experimental platforms, entering more and more mass-produced cars.

Then, a compact tutorial of the lidar-based perception systems for autonomous driving was presented. Three levels of

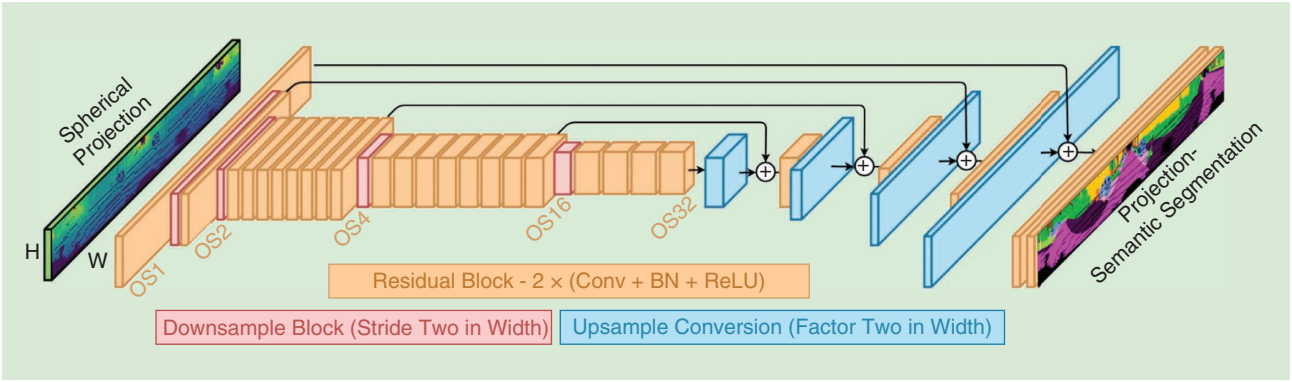

FIGURE 7. The neural network structure proposed in RangeNet [51]. Conv: conventional; ReLU: rectified linear unit; BN: batch normalization.

information providing by-perception systems along with typical processing pipelines were introduced. Generally, compared to cameras or radar, lidar is the most precise sensor used to measure range; therefore, the physical information (the objects' positions, headings, shapes, and so on) evaluated by lidar-based algorithms is highly reliable. However, semantic description is the shortcoming of lidar; this is caused by the lidar's poor resolution and its essence as a distance-measuring sensor, not a contextual sensor. Its fusion with cameras remedies lidar's weakness in recognition.

The intention-prediction level is independent of specific sensors, although it is strengthened by the precise physical information brought by lidar. Applying DL in lidar's 3D data will be one of the most important directions in the future. Lacking a huge number of annotated 3D point-cloud data sets was the bottleneck for successfully applying DL methods; however, things are changing. The aforementioned SemanticKITTI initiated a good start, and the results achieved by RangeNet++ are quite impressive. From our point of view, the algorithms used for extracting more accurate physical information and squeezing lidar's potential in semantic estimation are future directions. And, of course, with the swift progress of new lidars, new algorithms adapted to specific lidars will emerge.

Authors

You Li (you.li@renault.com) received his B.S. degree from Nanjing University of Science and Technology, China, in 2006, and his M.S. degree from Shanghai Jiao Tong University, Shanghai, China, in 2010. He completed his Ph.D. thesis on machine perception applied to intelligent vehicles at the Université de Technologie de Belfort Montbéliard, France, in 2013. Since 2017, he has been a senior research engineer in the research division of Groupe RENAULT, Guyancourt, France. Since 2014, he has worked as a postdoctoral researcher at the French National Center for Scientific Research, studying unmanned surface vehicles. In 2016, he joined The Netherlands Organisation for Applied Scientific Research as a research scientist working on sensor fusion for autonomous vehicles. His main research interest is lidar-based perception systems for autonomous vehicles.

Javier Ibanez-Guzman (javier.ibanez-guzman@renault.com) received his M.S.E.E. degree from the University of Pennsylvania, Philadelphia, as a Fulbright Scholar and his Ph.D. degree from the University of Reading, Pennsylvania, on an SERC-UK fellowship. In 2011, he was a visiting scholar at the University of California, Berkeley (CITRIS), working on connected vehicle applications. He is currently a corporate expert in autonomous systems at Groupe RENAULT, Guyancourt, France, as well as a member of the technical staff at Renault S.A., working on autonomous vehicle navigation technologies and driving assistance systems. Previously, he was a senior scientist at a national research institute in Singapore, where he spearheaded work on autonomous ground vehicles operating in unstructured environments. He has several publications and patents in the robotics and automotive domains and has successfully supervised several Ph.D. students. He is a Chartered Engineer and a fellow of the Institute of Engineering Technology. He is a Member of the IEEE.

References

[1] J. L. Leonard, "A perception-driven autonomous urban vehicle," *J. Field Robot.*, vol. 25, no. 10, pp. 727–774, 2008. doi: 10.1002/rob.20262.

[2] National Research Council. *Laser Radar: Progress and Opportunities in Active Electro-Optical Sensing*. Washington, D.C.: National Academies Press, 2014.

[3] C. Urmson, A. Joshua, B. Drew, B. Christopher, B. Robert, M. N. Clark, J. Dolan, D. Duggins et al., "Autonomous driving in urban environments: Boss and the urban challenge," *J. Field Robot.*, vol. 25, no. 8, pp. 425–466, 2008.

[4] A. Petrovskaya and S. Thrun, "Model based vehicle detection and tracking for autonomous urban driving," *Auton. Robots*, vol. 26, nos. 2–3, pp. 123–139, 2009. doi: 10.1007/s10514-009-9115-1.

[5] Y. Li, "Stereo vision and LIDAR based dynamic occupancy grid mapping: Application to scenes analysis for intelligent vehicles," Ph.D. dissertation, Univ. Technol. Belfort-Montbeliard, 2013.

[6] U. Wandinger, "Introduction to LiDAR," in *Springer Series in Optical Sciences*, C. Weitkamp Ed. New York: Springer-Verlag, vol. 102, 2005.

[7] *Interpretation sheet 1-Safety of laser products—Part 1: Equipment classification and requirements*. IEC Standard 60825-1, 2017.

[8] D. F. Pierrottet, F. Amzajerdian, L. Petway, B. Barnes, G. Lockard, and M. Rubioand, "Linear FMCW laser radar for precision range and vector velocity measurements," *MRS Proc.*, vol. 1076, 2008.

[9] S. Piatek, "Lidar and other techniques," Hamamatsu Photonics, Japan, Tech. Rep., 2017. [Online]. Available: https://www.hamamatsu.com/sp/hc/osh/lidar_webinar_12.6.17.pdf

[10] R. Horaud, M. Hansard, G. Evangelidis, and C. Ménier, "An overview of depth cameras and range scanners based on time-of-flight technologies," *Mach. Vis. Appl.*, vol. 27, no. 7, pp. 1005–1020, 2016. doi: 10.1007/s00138-016-0784-4.

[11] J. O'Neill, W. T. Moore, K. Williams, and R. Bruce, "Scanning system for LiDAR," U.S. Patent 8 072 663B2, 2011.

[12] H. W. Yoo and G. Schitter, "MEMS-based LiDar for autonomous driving," *Elektrotech. Inftech.*, vol. 135, no. 6, pp. 408–415, 2018. doi: 10.1007/s00502-018-0635-2.

[13] N. Druml, I. Maksymova, T. Thomas, D. Lierop, M. Hennecke, and A. Foroutan, "1D MEMS micro-scanning LiDar," in *Proc. 9th Int. Conf. Sensor Device Technologies and Applications*, Venice, Italy, 2018.

[14] F. Amzajerdian, V. E. Roback, A. Bulyshev, P. F. Brewster, and G. D. Hines, "Imaging flash LiDar for autonomous safe landing and spacecraft proximity operation," in *Proc. AIAA SPACE*, Long Beach, CA, 2016. doi: 10.2514/6.2016-5591.

[15] C. V. Poulton, M. J. Byrd, E. Timurdogan, P. Russo, D. Vermeulen and M. R. Watts, "Optical phased arrays for integrated beam steering," in *Proc. 2018 IEEE 15th Int. Conf. Group IV Photonics (GFP)*, pp. 1–2. doi: 10.1109/GROUP4.2018.8478729.

[16] P. F. McManamon, T. A. Dorschner, D. L. Corkum, L. J. Friedman, D. S. Hobbs, M. Holz, S. Liberman, H. Q. Nguyen et al., "Optical phased array technology," *Proc. IEEE*, vol. 84, no. 2, pp. 268–298, 1996. doi: 10.1109/5.482231.

[17] I. Takai, H. Matsubara, M. Soga, M. Ohta, M. Ogawa, and T. Yamashita, "Single-photon avalanche diode with enhanced NIR-sensitivity for automotive LIDAR systems," *Sensors (Basel)*, vol. 16, no. 4, pp. 459–468, 2016. doi: 10.3390/s16040459.

[18] S. Gnecchi and C. Jackson, "A 1×16 SiPM array for automotive 3D imaging LiDAR systems," in *Proc. Int. Image Sensor Workshop (IISW)*, Hiroshima, Japan, 2017, pp. 133–136.

[19] S. Michaud, J.-F. Lalonde, and P. Giguère, "Towards characterizing the behavior of LiDARs in snowy conditions," in *Proc. IEEE/RSJ IROS Workshop on Planning, Perception and Navigation for Intelligent Vehicles*, 2015.

[20] M. Kutila, P. Pyykonen, H. Holzhuter, M. Colomb, and P. Duthon, "Automotive LiDAR performance verification in fog and rain," in *Proc. Int. Image Sensor Workshop (IISW)*, pp. 1695–1701. doi: 10.1109/ITSC.2018.8569624.

[21] M. Himmelsbach, A. Müller, T. Lüttel, and H.-J. Wünsche, "LIDAR based 3D object perception," in *Proc. 1st Int. Workshop on Cognition for Technical Systems*, 2008.

[22] I. Bogoslavskyi and C. Stachniss, "Fast range image-based segmentation of sparse 3D laser scans for online operation," in *Proc. 2016 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, pp. 163–169. doi: 10.1109/IROS.2016.7759050.

[23] D. Zermas, I. Izzat, and N. Papanikolopoulos, "Fast segmentation of 3D point clouds: A paradigm on LiDAR data for autonomous vehicle applications," in *Proc. 2017 IEEE Int. Conf. Robotics and Automation (ICRA)*, Singapore, pp. 5067–5073. doi: 10.1109/ICRA.2017.7989591.

[24] T. Chen, B. Dai, D. Liu, and J. Song, "Performance of global descriptors for velodyne-based urban object recognition," in *Proc. 2014 IEEE Intelligent Vehicles Symp.*, pp. 667–673. doi: 10.1109/IVS.2014.6856425.

[25] C. Premebida, O. Ludwig, and U. Nunes, "Exploiting LIDAR-based features on pedestrian detection in urban scenarios," in *Proc. 2009 IEEE Int. Conf. Intelligent Transportation Systems (ITSC)*, pp. 1–6. doi: 10.1109/ITSC.2009.5309697.

[26] B. Douillard, J. Underwood, V. Vlaskine, A. Quadros, and S. Singh, "A pipeline for the segmentation and classification of 3D point clouds," in *Experimental Robotics (Springer Tracts in Advanced Robotics)*, O. Khatib, V. Kumar, and G. Sukhatme, Eds. Berlin, Heidelberg: Springer-Verlag, 2014.

[27] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 5, pp. 433–449, 1999. doi: 10.1109/34.765655.

[28] D. Z. Wang, I. Posner, and P. Newman, "What could move? Finding cars, pedestrians and bicyclists in 3D laser data," in *Proc. 2012 IEEE Int. Conf. Robotics and Automation (ICRA)*, pp. 4038–4044. doi: 10.1109/ICRA.2012.6224734.

[29] Y. Li and Y. Ruichek, "Moving objects detection and recognition using sparse spacial information in urban environments," in *Proc. 2012 IEEE Intelligent Vehicles Symp.*, pp. 1060–1065. doi: 10.1109/IVS.2012.6232205.

[30] E. Capellier, F. Davoine, V. Cherfaoui, and Y. Li, "Evidential deep learning for arbitrary LIDAR object classification in the context of autonomous driving," in *Proc. 2019 IEEE Intelligent Vehicles Symp.*, pp. 1304–1311. doi: 10.1109/IVS.2019.8813846.

[31] S. Challa, M. R. Morelande, D. Mušicki, and R. J. Evans, *Fundamentals of Object Tracking*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[32] F. Moosmann and C. Stiller, "Joint self-localization and tracking of generic objects in 3D range data," in *Proc. 2013 IEEE Int. Conf. Robotics and Automation (ICRA)*, pp. 1146–1152. doi: 10.1109/ICRA.2013.6630716.

[33] H. Cho, Y. Seo, B. V. K. V. Kumar, and R. R. Rajkumar, "A multi-sensor fusion system for moving object detection and tracking in urban driving environments," in *Proc. 2014 IEEE Int. Conf. Robotics and Automation (ICRA)*, pp. 1836–1843. doi: 10.1109/ICRA.2014.6907100.

[34] A. A. Rachman, "3D-LIDAR multi object tracking for autonomous driving," Master's thesis, Delft Univ. Technol., Delft, The Netherlands, 2017.

[35] X. Zhang, W. Xu, C. Dong, and J. M. Dolan, "Efficient L-shape fitting for vehicle detection using laser scanners," in *Proc. 2017 IEEE Intelligent Vehicles Symp. (IV)*, pp. 54–59. doi: 10.1109/IVS.2017.7995698.

[36] T. Johansson, "LiDAR clustering and shape extraction for automotive applications," Master's thesis, Chalmers Univ. Technol., Gothenburg, Sweden, 2017.

[37] S. Kraemer, C. Stiller, and M. E. Bouzouraa, "Lidar based object tracking and shape estimation using polylines and free-space information," in *Proc. 2018 IIEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, pp. 4515–4522. doi: 10.1109/IROS.2018.8593385.

[38] C. Miyajima and K. Takeda, "Driver-behavior modeling using on-road driving data: A new application for behavior signal processing," *IEEE Signal Process. Mag.*, vol. 33, no. 6, pp. 14–21, 2016. doi: 10.1109/MSP.2016.2602377.

[39] N. Deo, A. Rangesh, and M. M. Trivedi, "How would surround vehicles move? A unified framework for maneuver classification and motion prediction," *IEEE Trans. Intell. Veh.*, vol. 3, no. 2, pp. 129–140, 2018. doi: 10.1109/TIV.2018.2804159.

[40] D. J. Phillips, T. A. Wheeler, and M. J. Kochenderfer, "Generalizable intention prediction of human drivers at intersections," in *Proc. 2017 IEEE Intelligent Vehicles Symp. (IV)*, pp. 1665–1670. doi: 10.1109/IVS.2017.7995948.

[41] N. Deo and M. M. Trivedi, "Multi-modal Trajectory Prediction of Surrounding Vehicles with Maneuver based LSTMs," in *Proc. 2018 IEEE Intelligent Vehicles Symp. (IV)*, pp. 1179–1184. doi: 10.1109/IVS.2018.8500493.

[42] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social LSTM: Human trajectory prediction in crowded spaces," in *Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 961–971. doi: 10.1109/CVPR.2016.110.

[43] N. Deo and M. M. Trivedi, "Convolutional social pooling for vehicle trajectory prediction," in *Proc. 2018 IEEE/CVF Conf. Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1468–1476. doi: 10.1109/CVPRW.2018.00196.

[44] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, and F. Nashashibi, "Non-local social pooling for vehicle trajectory prediction," in *Proc. 2019 IEEE Intelligent Vehicles Symp. (IV)*, pp. 975–980. doi: 10.1109/IVS.2019.8813829.

[45] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.

[46] M. Velas, M. Spanel, M. Hradis, and A. Herout, "CNN for very fast ground segmentation in velodyne LiDAR data," in *Proc. 2018 IEEE Int. Conf. Autonomous Robot Systems and Competitions (ICARSC)*, pp. 97–103. doi: 10.1109/ICARSC.2018.8374167.

[47] B. Yang, W. Luo, and R. Urtasun, "Pixor: Real-time 3D object detection from point clouds," in *Proc. 2018 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 7652–7660. doi: 10.1109/CVPR.2018.00798.

[48] X. Chen, H. Ma, J. Wan, B. Li, and T. Xia, "Multi-view 3D object detection network for autonomous driving," in *Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 6526–6534. doi: 10.1109/CVPR.2017.691.

[49] E. Capellier, F. Davoine, V. Cherfaoui, and Y. Li, "Transformation-adversarial network for road detection in LIDAR rings, and model-free evidential road grid mapping," in *Proc. 11th PPNIV - IROS Workshop on Planning, Perception, Navigation for Intelligent Vehicle*, 2019.

[50] D. Frossard and R. Urtasun, "End-to-end learning of multi-sensor 3D tracking by detection," in *Proc. 2018 IEEE Int. Conf. Robotics and Automation (ICRA)*, pp. 635–642. doi: 10.1109/ICRA.2018.8462884.

[51] A. Milioto, I. Vizzo, J. Behley, and C. Stachniss, "RangeNet++: Fast and accurate LiDAR semantic segmentation," in *Proc. 2019 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, pp. 4213–4220. doi: 10.1109/IROS40897.2019.8967762.

[52] R. Q. Charles, H. Su, M. Kaichun, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3D classification and segmentation," in *Proc. 2017 IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, pp. 77–85. doi: 10.1109/CVPR.2017.16.

[53] B. Wu, A. Wan, X. Yue, and K. Keutzer, "SqueezeSeg: Convolutional neural nets with recurrent CRF for real-time road-object segmentation from 3D LiDAR point cloud," in *Proc. 2018 IEEE Int. Conf. Robotics and Automation (ICRA)*, pp. 1887–1893. doi: 10.1109/ICRA.2018.8462926.

[54] J. Behley, M. Garbade, A. Milioto, J. Quenzel, S. Behnke, C. Stachniss, and J. Gall, "SemanticKITTI: A dataset for semantic scene understanding of LiDAR sequences," in *Proc. IEEE/CVF Int. Conf. Computer Vision (ICCV)*, 2019, pp. 9297–9307.

[55] "HDL-64-E," Velodyne Lidar, San Jose, CA. [Online]. Available: https://www.velodynelidar.com/hdl-64e.html

[56] "EVAPS," Mov'eo, Rouen. [Online]. Available: https://pole-moveo.org/projets/evaps/

[57] "Valeo Scala," Valeo, Paris. [Online]. Available: https://www.valeo.com/en/valeo-scala/

[58] A. Pacala, "How multi-beam flash lidar works," Ouster, San Francisco, Nov. 8, 2018. [Online]. Available: https://www.ouster.io/blog-posts/2018/11/8/how-multi-beam-flash-lidar-works

[59] Youtube, "GeigerCruizer demo system images to 300 meters." [Online]. Available: https://www.youtube.com/watch?v=K2YkrFSJc3A

[60] The KITTI Vision Benchmark Suite. [Online]. Available: http://www.cvlibs.net/datasets/kitti/

[61] The KITTI Vision Benchmark Suite, "Bird's eye view evaluation 2017." [Online]. Available: http://www.cvlibs.net/datasets/kitti/eval_object.php?obj_benchmark=bev

[62] Semantic KITTI. [Online]. Available: http://www.semantic-kitti.org/

SP