Département de Génie Électrique & Informatique

4th year - Automatic Electronic - Embedded Systems

PIR : State of the art

Embedded technologies in an autonomous car for perception and detection of the environment

May 30, 2020

Thomas Zennaro, *Student, INSA Toulouse,*
Romain Gary, *Student, INSA Toulouse,*
Alexandre Dimnet, *Student, INSA Toulouse,*
Xavier Bourlot, *Student, INSA Toulouse,*
Pauline Combes, *Student, INSA Toulouse,*
Thibault Jean, *Student, INSA Toulouse,*

Tutors :

M. Patrick TOUNSI
M. Sébastien DI MERCURIO

***Abstract*—Autonomous cars have an architecture that is still poorly standardized, the system and software architectures presented in this document show certain similarities. It is the same thing for the methods of perception of the environment, there are many technologies each having their advantages and their defects. The optimization of the performances thus inevitably passes by a coupling of these technologies, in order to make the system more robust, safer and reliable. The weaknesses of these systems are at the heart of research in the automotive industry. Through this document we review what exists and innovations for future vehicles on our roads.**

***Index Terms*—Autonomous car, Perception, Lidar, Radar, 3D Detection**

Introduction

SINCE many years, the autonomous vehicle is a project which inspires several research centers and car makers. Indeed, we can take the example of the Google Car which showed satisfactory results for autonomous driving with its 500000 kilometers on the meter [1]. Furthermore, according to [1], the IEEE[1] asserts that 75% of the vehicles which will circulate on the American roads in 2040 will be fully autonomous.

Faced with this desire to change our travel habits, it is crucial to master all the embedded technologies that we implement. The safety of people is one of the main stakes especially when we know that in France, there is a death for 12.5 million kilometres travelled, and that humans are liable for 80% of road accidents [2]. This will be based on many factors. Both [1] and [2] insist on the reliability of the different sensors, but also the ability of the vehicle's artificial intelligence to take into account the different disturbances in its environment in order to make the right decisions at any time. These are part of the technological challenges of tomorrow.

In this document, we start by talking about the characteristics of the current autonomous car and the general system architecture. Then, we present the main technologies used by autonomous vehicles to collect the necessary data from its environment to have the best possible perception. Finally, we explain some detection methods using these technologies in order to effectively acknowledge its surroundings.

[1]Institute of Electrical and Electronics Engineers

1. Autonomous car and Technological approaches

A. Self-driving car features

1) Brief definition of autonomous cars: In order to acquire this autonomy, the vehicle needs to collect information to know what it has to do at any given time. We can easily define an autonomous car as a vehicle which is sufficient for itself, that is to say that it has the ability to adapt to the various situations to which it is subjected by making a decision quickly and without human intervention [3].

According to both [3] and [4], there are five essential functions that the autonomous car must properly fill :

- *Perception* : need to analyse the surrounding environment.
- *Localization* : need to know a specific position.
- *Control* : apply a command to accelerate or to brake for example.
- *Planning* : select the control actions.
- *System management* : an artificial intelligence supervises the simultaneous functioning of embedded systems for the driving.

Fig. 1 illustrates these functions in a simplified scheme.

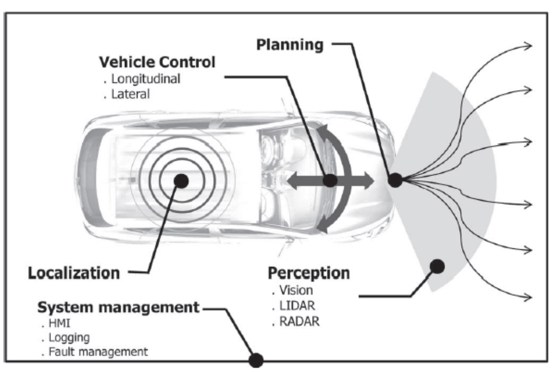

Fig. 1: Conceptual scheme of the main functions for an autonomous car [4]

2) Elaborated system architectures: There is no doubt that to fill these requirements, we need to define functional components to integrate inside the car which will be the main parts of the processing computer units. For that, there exist two types of architectures : the centralized system and the distributed system.

a) Centralized architecture: According to [4], single computing unit is necessary to connect every component such as actuators and sensors, as shown in Fig. 2. So, this architecture has the main advantage to simplify any configurations with the car devices, communication and information exchanges. Indeed, we have directly access to all the data coming from the sensors and actuators and as a result processing information quickly.

However, there is a challenge towards the implementation of the driving algorithms due to the significant amount of data received, which makes integration more complex and this has an impact in the system reliability [4], particularly if a bug appears and disrupts the functioning of the car.

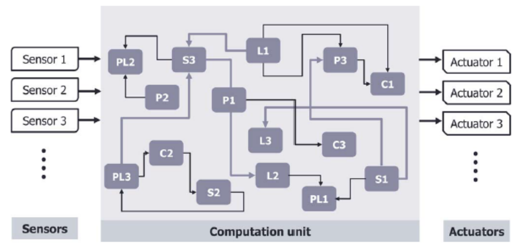

Fig. 2: Schematic principle of a centralized system [4]

b) Distributed architecture: For a distributed architecture, we can find two or more computing units, thus it allows to decentralized into local computing units whose role is to monitor and deal with either a part of the sensors or actuators. A conceptual model is shown by Fig. 3.
As said in the article [4], we can reduce the complexity of data process, therefore the used algorithms may be independently developed and easier to implement inside the driving system, contrary to the case of centralized system architecture. The concurrent computation is otherwise possible, which guarantee a better performance of the whole system. Furthermore, a distributed system architecture will be safer and more robust towards dysfunctions thanks to the independence between all modules and so, the faulty module becomes easier to detect and less costly to patch. The maintainability of both software and sensor configurations is improve [4].
As we will see later, this judicious distribution will be beneficial for optimal treatments when detecting objects and forms, for instance.

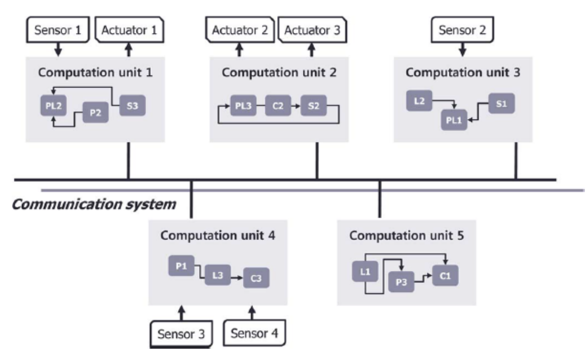

Fig. 3: Schematic principle of a distributed system [4]

B. *Connectivity and software architectures of autonomous driving systems*

In this subsection, we will focus on the different type of architectures present in the literature.

1) Connectivity of autonomous driving system: First of all, [5] exposes two ways to implement an autonomous driving system.

a) Ego-only Systems: The most common way to implement this type of system is to design a software that is not dependent of the other cars on the road. Therefore, the car will use only its own information provided by its own sensors to achieve an autonomous driving. This kind of methods are called "Ego-only system". According to [5], this type of system is the most developed now, because it is easier to implement a such system than a connected system.

b) Connected systems: Contrary to ego-only systems, the connected systems base the autonomous driving on shared information between the cars driving on the road. This type of system is not really developed yet, because it implies a big amount of autonomous car on the road. To communicate, the cars can use two protocols : the IP[2] protocol or the ICN[3]. According to [5], the IP protocol is really difficult to implement in the context of connected autonomous car since the connection between cars is unstable and the amount of data exchanged is huge. An other issue with connected autonomous car is the mobility of the vehicles.
The idea of the ICN is to send the data into an area collecting the data instead of an address. (The IP protocol is an address to address protocol). However, a lot of challenges are related to these methods as the huge amount of data, how to use consistently the data, etc. Because of the complexity of the problem, it does not exist yet an operational connected system; that is why, in this paper, we will focus on Ego-only System.

2) Software architectures for autonomous driving systems: According to [5], we can find two type of software architectures in the literature : modular systems and end-to-end driving systems.

a) Modular Systems: A modular system denotes a software using several modules to achieve the autonomous driving. Generally, we can identify four steps between sensors and actuators : Object detection and tracking, Behaviour prediction, planning and control. [6] gives us an example of a such system, presented in Fig 4. In this figure, we can well identify the object detection and tracking blocks in blue, aiming to detect where is the road, if there are pedestrians on the road, etc. All these detection techniques are developed in section 3. Second, the goal of the SLAM module and the decision module is to take a decision from the precedent step, it constitutes the behaviour predicting module. The other green blocks aim to plan the trajectory to chose, it is the planning module. Finally, the orange blocks constitute the control module. Hence, this type of architectures uses pedestrian detection, road and lanes detection and magnetic detection, respectively explained in section 3.A, 3.B, 3.C.

[2]Internet protocol
[3]Information-centring networking

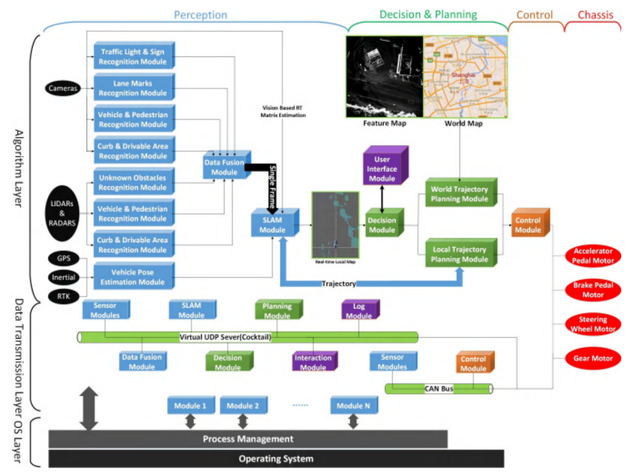

Fig. 4: Example of a Modular software architecture [6]

b) End-to-end driving architectures: As explained in [5], the end-to-end driving architectures associate directly the data collected from the sensors with actions to do. The most common way to implement such architectures is to use deep learning methods. These methods are developed in section 3.D. Finally, [5] shows a schematic summary of the differences between software architectures (see Fig 5).

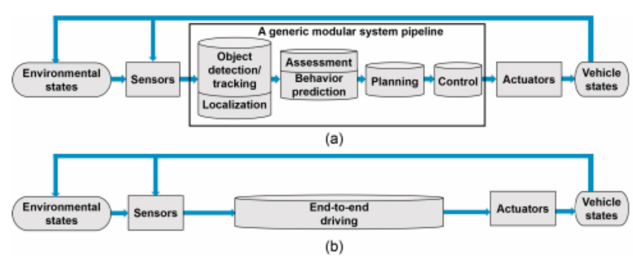

Fig. 5: Difference between modular and end-to-end driving architectures [5]

2. Components for perception purposes of the environment

A multitude of equipment is added to a car to guarantee its autonomous management. We present the main sensors which each have a very special mission and which are now widespread in the automotive industry.

A. Lidar

1) Main features: Lidar[4] is a high-tech sensor based on a method of transmission and retrieval of information by laser. It can be caricatured as the vehicle's vision organ. Lidar sensors can work in several dimensions as precised in [7]. At the mono-dimensional level, they act as point distance measurement systems. For that, they are directed on a particular target. By varying the measurement beam in a plane, precise information on the distance and angle of the vehicle relative to an object is obtained, as said in both [7] and [5]. Data capture is made at regular time intervals. Finally, for a 3D view of space, it is necessary to move several transmitting and receiving systems, therefore to have several scanners in order to reconstruct a target in space as accurately as possible. As said in [8], reconstructed image quality is highly dependent on horizontal and vertical parameter calibration and of the resolution.

Lidar sensors applied in automotive area are based on TOF[5] principle. This technique consists in measuring the time taken by a laser pulse from a transmitter to travel back and forth between a sensor and an object, which then allows to compute the precise distance between the sensor source and the selected target. The TOF model is shown by Fig. 6 : a *Light source* generates the laser beam while a *photosensor* is ready to receive the reflection of the resulting laser and converts it into a usable electrical signal which will then be amplified to be able to process the information easily. At this moment, a *timer* stops counting and then we collect the travel time [7].

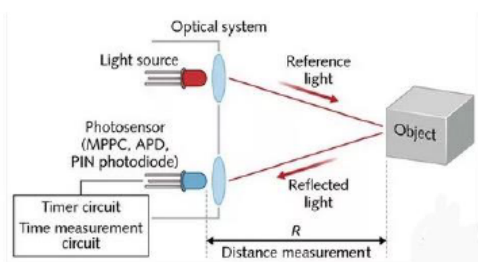

Fig. 6: Principle schematic of TOF ranging [7]

A study related in [8] determined an average of Lidar detection range of several types of object and pedestrians that we can often met on the road. This is shown in Table. I. These values may vary significantly depending on observations conditions and also if the target is moving.

We can noticed that highly reflective objects such as other vehicles are detected further unlike pedestrians or road lines. This is normal since Lidar is sensitive to reflection phenomena and for instance, a truck is a perfect candidate because its bodywork is usually covered with metal. Taking into account the adaptability of the laser transmitter, the energy of the laser pulse and the sensitivity of the receiver, Lidar sensors can reliably measure objects at long distances and with low reflectance. So a cyclist, for example, might have the chance to be spotted earlier.

Target name	**Range detection** (m)
Pedestrian (side view)	59
Bicycle (side view)	45
Car	232
Truck	237
Road sign	142

TABLE I: *Table relating maximal range detection of some 3D objects and pedestrians by Lidar (data found in [8])*

2) Lidar types: The article [7] inform us that there exist two Lidar types : scanning and non-scanning Lidar.

a) Scanning Lidar: Typically, scanning Lidar can be diode laser which scans in rotation. We meet single-line scanning Lidar and multi-line scanning Lidar.

The first type is a laser which operates at high frequency combined with a rotation scanner. It is used in two-dimensional image capture. As noticed in [7] many advantages such as low power consumption and small size gives it a place among the sensors used in a autonomous car. Furthermore, the single-line scanning Lidar is often used as obstacle detector [7].

Then, a multi-line scanning Lidar combines an angular scanning system with mirrors that deflect the emission and reception axis of a laser beam [7]. We are able to locate the echoes collected in 3D environment. This allows the target to be mapped in 3D with accuracy and range over several hundred meters, which is necessary depending on the speed when the vehicle is moving. It also allows an accuracy SLAM[6] [7].

b) Non Scanning Lidar: The 3D Flash Lidar is the most widespread among non scanning Lidar sensors. As previously, a laser is emitted to the target but this time, a transmission optical system is going to expand the emission range by creating a deflection in several directions of the light beam from the source. On the contrary, during the reception of the light reflection, an optical system will focus incoming beams towards an array detector [7]. After data capture, a process circuit reconstructs a 3D pixel-by-pixel image taking into account the amplitude of the received beams and their origin.

3) Disturbances due to the weather: Both [5] and [8] agree that weather conditions have a significant influence on the reliability of Lidar sensors. Indeed, studies made

[4]Light Detection and Ranging
[5]Time Of Flight
[6]Simultaneous Localization and Mapping

in [8] during periods of fog and rain showed that laser intensity of Lidar sensors are attenuated : atmospheric moisture on the air and on roads severely impacts sensors' performance. Furthermore, as for human, the visibility on the road also decreases. Therefore, distance detection is reduced. For instance, with a clear weather, visibility is close to $5km$ whereas during dense fog, it goes down to $40m$ [8].

B. Radar

1) Main features and principle: Radar, as explained in [9] and [6], is a sensor based on waves and echoing effect. Indeed, the RADAR send a wave from an antenna and receive a few nanoseconds later the echo of this wave. Then, by analysing this signal, the sensor is able to draw the shape of the environment. The advantage given by this technology, according to [9], is the fact that one emission return information on multiple targets. Moreover, contrary to the Lidar technology, the RADAR works well in difficult conditions such as the rain, the fog and other meteorological conditions.
In addition, the reliability of the RADAR technology decrease with its range, as said in [6]. Therefore, the RADAR is mainly used to detect range and distance in autonomous car.
Finally, the technology used in autonomous car is generally millimeter Wave Radar, but we will see further that this technology can be improved by signal processing or additional features.

2) Technologies of millimeter RADAR sensors: In this section, we will study more in details how works the millimeter radar technology and how we can improve the resolution and the reliability of this system.

a) Mode of operation of the millimeter RADAR technology: The article [9] explains how the millimeter RADAR technology is working.
As explained before, the RADAR send a beam of wave with a wavelength λ in the order of a few millimeters. When the wave touch an obstacle (it can be the ground, an object on the road, etc.), a part of it comes back to the sensor by echo effect. if we now look at the signal returned by the echo, we observe than depending on the angle, the aperture of the sensor and the distance of the obstacle, the signal have a different spectrum, as shown in Fig. 7.

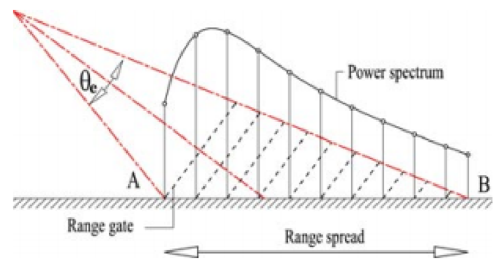

Fig. 7: Expected power return with RADAR technology [9]

Furthermore, the document [9] gives us the equation of the beam width θ_e in function of the wavelength λ and the antenna aperture D :

$$\theta_e = 1.02\frac{\lambda}{D} \tag{1}$$

Hence, increasing the aperture will decrease the beam width and then improve the accuracy of the radar. However, in a context of an autonomous car, we have a limited antenna aperture since we have a limited antenna size. Therefore, the stake of the RADAR technology in autonomous car is to find a compromise between the limited size of the antenna and the accuracy of the results. Moreover, a lack of accuracy is caused by the noise present in the environment of the RADAR. Since an autonomous car evolves in an outdoor environment which is particularly noisy, the accuracy of the RADAR system is affected. For these two reasons (the noise and the size of the antenna), the RADAR technology cannot be used alone, it has to be completed and improved by signal processing or additional sensors and features.

b) Improving RADAR technology by signal processing algorithm : Radar Ground segmentation system: As developed in [9], the RADAR technology cannot do the difference between the ground and an obstacle that is not part of the ground. Hence, the idea of this process is to anticipate the wave returned by the ground and to erase it from the signal to exclusively see the obstacles on the road. Briefly summarized, it erases the noise due to the ground.
The first step in the echo segmentation process is to provide a model of the echo resulting from the ground.This model depends of course of the chosen beam width, the azimuth angle of the sensor and the tilt of the vehicle. This step is called extraction of the background. For example, the expected model of a ground echo is given in Fig. 8.

Then, the second step of the algorithm is to sort the echos resulting from the ground and the one resulting from an other obstacle. For that, the system compares the power spectrum and detects the peaks corresponding to a ground response, as shown in Fig. 9. Let's note that in parallel, a CUSUM[7] algorithm is achieved to adapt the threshold level of the obstacle detection with the field shape.

c) MIMO Radars: An other possibility to improve the accuracy of a millimeter wave RADAR is to provide

[7] Cumulative Sum

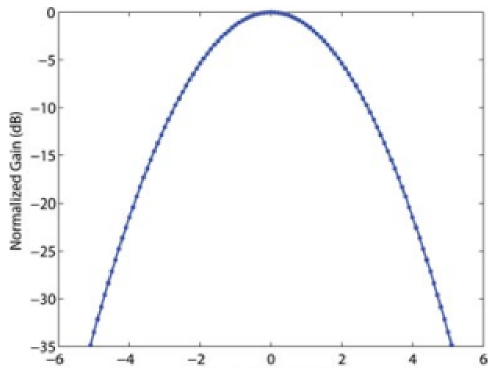

Fig. 8: Expected ground echo for a flat field and for an azimuth angle $\alpha = 32°$[9]

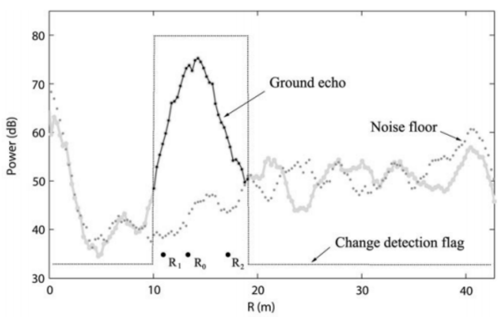

Fig. 9: Example of a detection of the ground echo [9]

several antennas. That the solution given by the MIMO[8] RADAR presented in [10]. The principle is quite simple, several antennas are placed in parallel to create an array of data collected by each antenna. Then, the system can precisely know the angle, the elevation and the distance of the obstacle by crossing the data of each antenna. For example, if an object is on the left side of the sensor, the left antennas will receive more power than the other ones and the system will be able to detect the object is on the left.

3) Limits of RADAR technology: The main limit concerning the use of the RADAR technology is the range reachable by the sensor. Indeed, as explained in [6], the range reachable by the sensor is directly connected to the resolution of the data collected. Therefore, to allow a better resolution, the RADAR technology can be used exclusively for a low or medium range detection.
An other limit is about the shape of the field. Indeed, as developed in [9], even with RADAR ground segmentation, if the shape of the field is too irregular, the system can interpret a slope as an obstacle to avoid.

[8]Multiple Input Miultiple Output

C. Stereovision / cameras

1) Main features and principle: The visual sensors most used in autonomous vehicle are monocular vision system, stereo vision system, and infrared vision system according [11] And [12]. Nowadays, monocular vision systems are widely used on some actual production vehicles.

a) Monocular Camera: The camera is defined as a passive sensor which returns raw optical data [5]. vision systems using cameras and stereo vision are very widely used in the perception of autonomous vehicles, as said in [11] and [5] its systems offer a greater variety of characteristics, for example, a greater angle of view, a color image and better resolution than LIDARs.

b) Stereovision: But they are less effective for depth measurements, this is why there are also stereo cameras that use two close cameras in order to reconstruct a 3D map of the environment as described [3] they remain still less reliable and have limited vision.

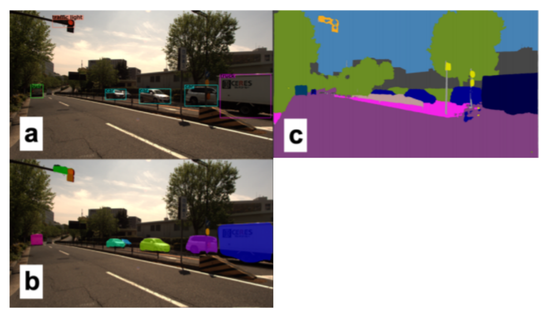

Fig. 10: Example of 3 different algorithms of images processing [5]

2) Image-based Object Detection: Monocular vision systems are used in the semantic segmentation of driving environment, target detection. The cameras can observe changes in the environment asynchronously, making them efficient for dynamic object detection while remaining inexpensive in memory and image processing. In Fig. 10 we can see three different detection algorithms "(a) A front facing camera's view, with bounding box results from YOLOv3 and (b) instance segmentation results from MaskRCNN. (c) Semantic segmentation masks produced by DeepLabv3"[5].This is an another advantage of this technology, the algorithms are numerous and well mastered (complex) due to the age of the technology.

3) Limits and solutions: The main problem according to [11] and [5] is that performance is very affected by changes in brightness, in season (snow, shadows). However, these systems being passive, they do not disturb active systems such as radars or LIDARs and can therefore easily be coupled to other perception systems.

3. Detection and perception methods

A. 3D perception of objects and pedestrians

1) 3D detection approach: Numerous methods allow to obtain only 2D representation of the space and obstacles around the autonomous car, but in order to detect accuracy shapes and to estimate the depth of objects, it is obvious that 3D detection becomes crucial. 3D detection is important for a realistic acknowledge of obstacles, but a main issue that we may notice is real-time operation. As said in both [5] and [13], the 3D Lidar can become a solution. Although they are rather sensitive towards fog and rain (*see 2.A.3*), 3D Lidar owns their lighting source thus they are more robust towards weather conditions than other sensors. Collected information is reliable and object localization is improve. As said by *Yurtsever and al.*[5], 3D Lidar algorithms, such as *VoxelNet*[9], use point cloud data which enable 3D visualization of objects and surfaces. Combined with several filtering techniques, we eliminate useless information and draw a landscape of the environment including both pedestrians and fixed objects. An overview is given by Fig. 11. As we can see in Fig. 11, the resolution can be further improved by associating 3D Lidar with cameras. Indeed, they specifically estimate the depth of objects. Moreover, radars are interesting for their ability in velocity estimation [5] and can be complementary for Lidar.
However, although this sensor is performing, 3D Lidar is expensive and that is why other methods are going to be preffered : the monocular-based 3D object detection is an example [13].

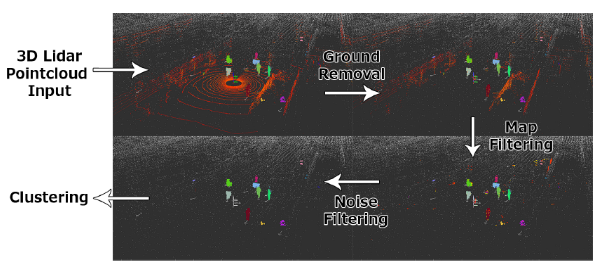

Fig. 11: Overview of pedestrians' detection with Lidar using point cloud data. [5]

2) Monocular-based 3D object detection: This method is inspired by image-based object detection whose purpose is to identify the size of a target and its positioning in a given space. [5]. As said by *Ma and al.*[13], this task is complex and its effectiveness is below what can produce the perception method with 3D Lidar. The weakness is as follows: the colored images obtained are mainly in 2D and do not accurately reflect reality, because there is a lack of depth information of the observed region.
The idea is to use monocular images with an improvement regarding the data input representation [13]. Furthermore, the approach of point clouds as used in 3D Lidar is preferred to reflect a 3D road environment. Fig. 12 shows the general overview of the method put in place. We can see two phases: 3D data generation and 3D Box estimation.

a) 3D data generation: This step consists of using CNN[10] [13] [14] in order to obtain 2D bounding area and depth information which are crucial for space reconstruction. Then through a camera it is possible to generate point cloud data based on gathered information. So in knowing the pixel coordinate (x', y') and the depth, we can easily find its matching (x, y, z) in 3D region [13]. Otherwise, RGB information can be added for each point. The equation system is as below : f is the focal length of the used camera, (C_x, C_y) the reference point coordinates and d is the image depth [13].

$$\begin{cases} x = (x' - C_x) * d/f \\ y = (y' - C_y) * d/f \\ z = d \end{cases} \tag{2}$$

b) 3D box estimation: In this step, the objective is to provide an accurate localization of the object [13]. For that, it is sufficient to proceed with data segmentation techniques [14], in order to calculate the volume inside each box (knowing height, width and depth (h, w, d)) obtained with CNN and then give an approximated value of the position of the target which will serve as a reference to know whether close points are in the same space plane of the target or not. The measure reference taken is the ground. Otherwise, segmentation techniques such as semantic segmentation allows to acknowledge the target class of belonging, whether it is a vehicle or a pedestrian, thus shapes can be easily identified [14]. Using architectures like DetNet[11] which implement 2D-3D algorithms, we finally find the mean position of a target.

c) Some results: The method proposed by *Ma and al.* [13] gives satisfactory results regardless of the difficulty level of detection, as seen in Table II. It comes first in terms of performance. An IoU of 0.5 is the minimum threshold for testing: higher is the IoU, that is, close to 1, greater is the overlap between estimation and reality. This means that we are approaching the exact position of the detected object.

[9]End to end network which extract point cloud data for 3D detection and generates rectangular bounding boxes

[10]Convolutional Neural Network : usually used in image recognition

[11]Deterministic Networking

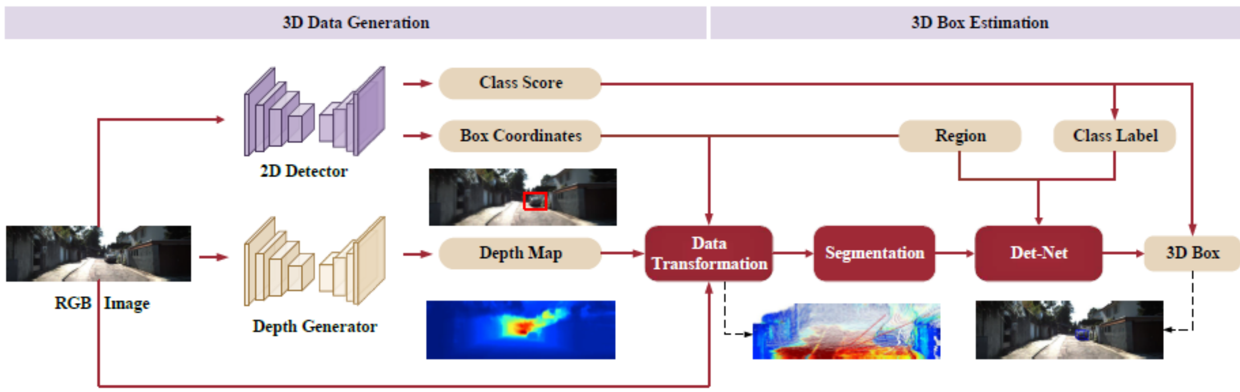

Fig. 12: Operating scheme of monocular-based detection

IoU[12]=0.5			
Method	Easy	Moderate	Hard
Deep3DBox	27.04	20.55	15.88
Multi-Fusion	47.88	29.48	26.44
MonoGRNet	50.51	36.97	30.82
Pseudo-Lidar	66.30	42.30	38.5
Monocular 3D suggested in [13]	68.86	49.19	42.24

TABLE II: *Table which evaluate the precision (%) for 3D detection of bounding boxes after some tests which 3 levels of difficulty. Comparison between Monocular-based 3D detection and other methods (data found in [13])*

B. Road and lanes detection

1) The using of the different sensors in road and lane detection; their benefits and their drawback:

a) Monocular Detection: Both [12] and [15] agree that the most commonly way to achieve road and lane detection is the monocular detection, i.e. using visual data from a camera. Indeed, since the lane marks are designed for human eyes, data from a visual sensor as a camera can easily be used for lane and road detection. Of course, as explained in [12], the camera needs enough resolution. This article shows that to be reliable, the minimum resolution needed has to associate at least one pixel to the lane marks. There are two main benefits with the monocular detection. First of all, as said before, it is the easiest way to detect the road thanks to basic algorithms and second, it is a cheap solution. Finally, the drawback of this solution is a bad reliability within difficult conditions such as weather perturbations (rain, fog, etc.) or light disturbances (overexposed image). Furthermore, a monocular detection system can be lost with a winding track.

b) LiDAR approaches: According to [12], a way to prevent the issues of weather and light disturbances is to use a LiDAR instead of a camera. Indeed, a LiDAR technology, as explained in section 2.A, is less sensible to weather conditions than a camera. However, the main drawback of this solution is the high cost of the sensor.

c) Stereo imaging: Finally, a compromise between price and reliability could be the stereo imaging, consisting in two cameras giving 3D information. A solution for road and lane detection based on stereo imaging method is given in [15]. This method keeps the depth information of the road, easing the identification of the curves and the noise due to the environment. The method exposed need three distinct steps to detect the lane presented as presented in Fig 13.

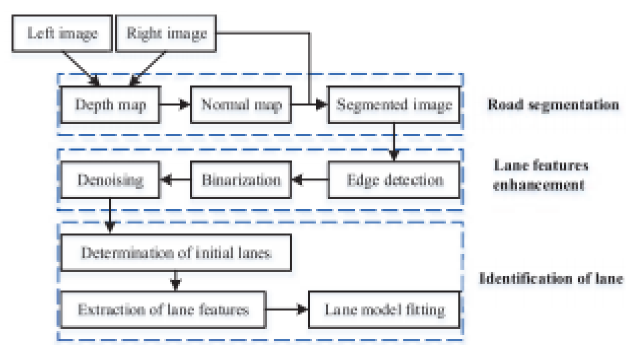

Fig. 13: Block diagram of a road detection solution based on stereo imaging [15]

The aim of the first step, called "Road segmentation" is to detect where is the road and to erase all the useless informations such as the roadside or the sky. The goal of the second step called "Lane features enhacement" is to detect and erase all the obstacles on the road that is not part of the road, i.e. the trucks, the cars and the other obstacles placed on the road. This is achieved by detecting the edges on the image and erasing the noise. Finally, the last step is to detect the lane with the useful information extracted from the two preceding steps. This is achieved thanks the

Hough transform and vanishing point algorithm. Finally, the study shows that a method based on stereo imaging and complex image processing algorithm can reach a reliability of 90%. Hence, the stereo imaging methods are quite reliable and constitute a compromise between price and reliability. However, a drawback of this kind of method is the high computing cost, leading to errors[12]. Moreover, even if it has a 90% reliability, a road detection system that can be used in real conditions needs a higher reliability, approaching 100%.

d) GPS and IMU[13] road detection: A good way to complement the preceding process to detect road and lanes is to use the information given by the GPS and the Inertial measurement Unit (IMU). Indeed, according to [12], the accuracy of the information given by the GPS on the car position is about 5 to 10 meters. Moreover, adding the information given by the IMU leads to an accuracy of 1m. Of course, this method does not allow an autonomous way to detect the road since the GPS signal can be easily lost. However, it can be used to compensate the errors given by stereo image techniques and to enhance the data given by the precedent exposed methods, especially when the conditions are difficult.

e) Drawbacks and benefits of the presented sensors: Finally, with the precedents sections, we can build a table to summarize the benefits and drawbacks of each sensors for the lane and road detection.

Sensor	Benefits	Drawbacks
Camera (monocular detection)	Easy to implement, cheap	Bad reliability with difficult conditions
LiDAR	Robustness especially with difficult conditions	High cost
Stereo imaging	Robustness, cheap	High computational cost
GPS and IMU	Accuracy up to 1 meter, works in difficult conditions	Risk to lose GPS signal

Fig. 14: Benefits and drawbacks of each sensor

2) System architectures for road and lane detection: In this subsection, we will focus on the solutions given at a system level for a reliable road and lane detection, using the knowledge from the precedent subsection. [12] exposes a generic block diagram for road and lane detection given in Fig 15. If we take a look on this block diagram, we can identify five blocks. Let's detail each goal of these functionalities :

a) Image Pre-processing: The image pre-processing block aims to detect and erase the parasite elements on the road such as the cars, the trucks or the roadsides. This feature has been explained before in section 3.B.1.c. It can use Camera (either monocular detection either stereo imaging depending on the wanted reliability) or LiDAR technology; but it can also use RADAR to detect obstacles like cars or trucks on the road.

[13]Inertial measurement Unit

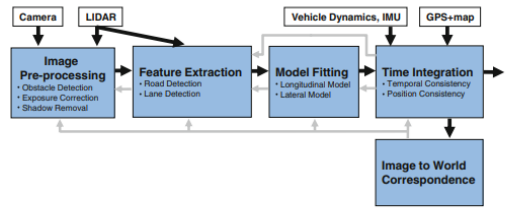

Fig. 15: Generic solution given by [12] for road and lane detection

b) Feature extraction: The goal of this second feature is to extract the lane position from the data given by the Image Pre-processing block. We can observe than the solution given in [15] use exclusively these two blocks to achieve road detection. Therefore, these features are the most important of the system and can lead to 90% of reliability. Hence, the aim of the other blocks is to improve this reliability by adding data from other sensors.

c) Model Fitting: The aim of this feature is to give a high level representation of the road and lanes found thanks to the precedent block.

d) Time integration: This module uses the GPS and IMU data to enhance and verify the hypothesis made by the precedent modules on the position of the road. Therefore, this feature will improve the reliability of the global system.

e) Image to world correspondence: Finally, if the time integration block indicates an error in the road detection provided by the three first modules, the Image To World matching feature will translate the detected position of the road to correct this mistake.

f) Summary and results: Finally, we can summarize the generic architecture for lane and road detection as following : A first detection is made by erasing the useless data of a visual or Lidar sensor. Then, the position found is compared to GPS and IMU data and corrected by the "Image to world correspondence" block.
If we now look at the results of this type of architecture, [12] indicates that some experiments have designed systems with a probability level of success with road and lane detection of 99%.

C. Deep learning method

As specified in [12] and [15], a quite recent breakthrough in self-driving car is the deep learning is the deep learning approach. Therefore, we will focus in this subsection on the recent progress in this field. We will first explain the principle of the deep learning method, then we will focus on how is it mostly used in a context of autonomous car and finally we will discuss about the results given by this method.

1) Principle of deep learning method: As explained in [16], the aim of this method is to use data from real human drivers in order to implement an auto-learning algorithm allowing the car to handle its acceleration. Thus, the car is supposed to be able to drive between two cars driven by humans, and adapting its speed by anticipating the acceleration or deceleration of the other drivers. For that, the deep learning method use a structure given in Fig 16. As we can observe, the deep learning method rely on two main function : the critic function and the policy function.

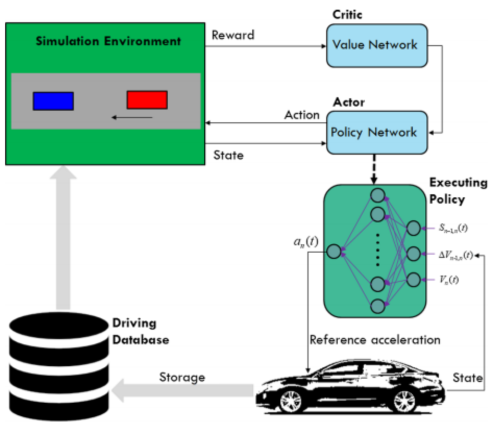

Fig. 16: Deep learning methods presented in [16]

The policy function aims to model the car-following comportment. Indeed, this function takes as input the relative speed between each vehicle, the desired speed, the follower speed and the gap distance between each vehicle and thanks to this information, provides the acceleration of the car as an output. To summarize, it is with this function that the car decides of the speed it needs to drive. It is the decision function.
The goal of the critic function is, as its name implies, to criticize the decision made by the car. This is achieved by comparing the result of the decision with the database. To create this database, as developed in [16], it is necessary to collect information from a large panel of human drivers in real conditions thanks to sensors. With this information, we have the information on "what an average human would have driven in this condition". Thanks to this data, the critic function can determine "Is the car acting like a real human?" and knows if the decision taken by the policy function was a human-likely decision or not.

2) Methods to determine the policy and critic function: In the literature, several ways to choose the policy and critic functions are developed, but the most promising one is the deep deterministic policy gradient (DDPG) since it ensures adaptability, generalization and accuracy [16]. Indeed, this type of method allows the adaptability of the policy function but also of the critic function. Both functions are represented with a network of states connected by arcs. The function will choose the state with the highest weight on the arc. For the critic function, at each iteration, the update of the weight of the arcs is made thanks to the comparison between expected acceleration and real acceleration (it gives the information : "Were the precedents path satisfying?"). With the difference between the precedent and the new critic function, and thanks to the actual state of the weight in the policy function, the weight of the arcs of the policy function are updated. We can then draw a summary of this process, showed in Fig 17. In this figure, θ represents the weights of the arcs in the critic network, μ the weights of the arcs of the policy function and $Q(s,a)$ the chosen critic function.

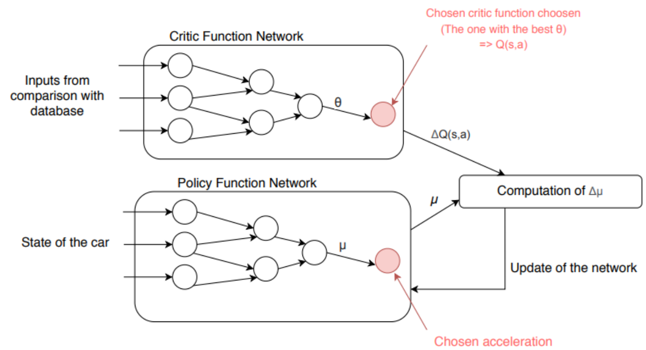

Fig. 17: Summary of the updating method for policy and critic function

3) Results given by deep learning methods: [16] presents some results of deep learning methods. With all the knowledge we have, the studies reach 18% regarding spacing requirements and 5% regarding speed requirement. Therefore, this kind of methods are, for now, not enough reliable to drive in real conditions, but these processes could be a serious lead for the self-driving car.

Conclusion

In this paper, we studied the recent breakthroughs and the actual work on the autonomous car. First of all, we have seen that real progresses has been achieved in sensor science, especially in LiDAR, RADAR and stereo-vision technologies. Thus, these new methods and the work realized in signal processing allow the sensors to reach a high level of reliability. Based on the data given by the sensors, several recent works studied the way to implement a safe and reliable driving system. We can see two main types of architecture in the literature : end-to-end architecture and modular architecture. We have seen through this paper than modular architectures are based on several modules such as road and lane detection or perception of objects and pedestrian. Real breakthroughs have been made on these subjects and the systems based on these modules are more and more reliable. However, it is not yet safe enough to drive in real conditions. Finally, the second type of architecture (end-to-end) is based on deep learning methods. We have seen that this type of methods are for now less reliable and safe than the modular methods. Nevertheless, ongoing works are bright and could lead to reliable and safe autonomous driving system in the future.

References

[1] R. MOONZUR, "Les voitures autonomes, bientôt sur nos routes ?" *Techniques de l'ingénieur*, Oct. 2012. [Online]. Available: https://www.techniques-ingenieur.fr/actualite/articles/les-voitures-autonomes-bientot-sur-nos-routes-1163/

[2] M. COMBE, "Quels freins au développement du véhicule autonome ?" *Techniques de l'ingénieur*, Jul. 2014. [Online]. Available: https://www.techniques-ingenieur.fr/actualite/articles/quels-freins-au-developpement-du-vehicule-autonome-737/

[3] B. C. ZANCHIN, R. ADAMSHUK, S. M. Mauro, and K. S. COLLAZOS, "On the instrumentation and classification of autonomous cars," *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 2631–2636, Oct. 2017. [Online]. Available: https://ieeexplore.ieee.org/document/8123022

[4] J. KICHUN, K. JUNSOO, D. Kim, J. CHULHOON, and S. MYOUNGHO, "Development of autonomous car—part i: Distributed system architecture and development process," *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, vol. 61, pp. 7131–7140, Dec. 2014. [Online]. Available: https://ieeexplore.ieee.org/document/6809196

[5] E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, "A survey of autonomous driving: Common practices and emerging technologies," *arXiv:1906.05113 [cs, eess]*, Jan 2020, arXiv: 1906.05113. [Online]. Available: http://arxiv.org/abs/1906.05113

[6] W. Zong, C. Zhang, Z. Wang, J. Zhu, and Q. Chen, "Architecture design and implementation of an autonomous vehicle," *IEEE Access*, vol. 6, pp. 21 956–21 970, 2018.

[7] J. Liu, Q. Sun, Z. Fan, and Y. Jia, "Tof lidar development in autonomous vehicle," in *2018 IEEE 3rd Optoelectronics Global Conference (OGC)*, Sep 2018, p. 185–190.

[8] M. HADJ-BACHIR and P. De Souza, "LIDAR sensor simulation in adverse weather condition for driving assistance development," Jan. 2019, working paper or preprint. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01998668

[9] G. Reina, J. Underwood, G. Brooker, and H. Durrant-Whyte, "Radar-based perception for autonomous outdoor vehicles," *Journal of Field Robotics*, vol. 28, no. 6, pp. 894–913, 2011.

[10] E. Miralles, T. Multerer, A. Ganis, B. Schoenlinner, U. Prechtel, A. Meusling, J. Mietzner, C. Weckerle, H. Esteban, M. Vossiek, M. Loghik, and V. Ziegler, "Multifunctional and compact 3d fmcw mimo radar system with rectangular array for medium-range applications," *IEEE Aerospace and Electronic Systems Magazine*, vol. 33, no. 4, pp. 46–54, 2018.

[11] Z. D. Y. C. C. Z. X. S. X. Z. J. X. W. S. Z. K. Yang DianGe, Jiang Kun, "Intelligent and connected vehicles: Current status and future perspectives," *Science China Technological Sciences*, vol. 61, pp. 1446–1471, 2018.

[12] A. Bar Hillel, R. Lerner, D. Levi, and G. Raz, "Recent progress in road and lane detection: a survey," *Machine Vision and Applications*, vol. 25, no. 3, pp. 727–745, 2014.

[13] X. Ma, Z. Wang, H. Li, P. Zhang, W. Ouyang, and X. Fan, "Accurate monocular 3d object detection via color-embedded 3d reconstruction for autonomous driving," 2019, p. 6851–6860. [Online]. Available: http://openaccess.thecvf.com/content_ICCV_2019/html/Ma_Accurate_Monocular_3D_Object_Detection_via_Color-Embedded_3D_Reconstruction_for_ICCV_2019_paper.html

[14] X. Chen, K. Kundu, Z. Zhang, H. Ma, S. Fidler, and R. Urtasun, "Monocular 3d object detection for autonomous driving," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2016, p. 2147–2156.

[15] C. Yuan, H. Chen, J. Liu, D. Zhu, and Y. Xu, "Robust lane detection for complicated road environment based on normal map," *IEEE Access*, vol. 6, pp. 49 679–49 689, 2018.

[16] M. Zhu, X. Wang, and Y. Wang, "Human-like autonomous car-following model with deep reinforcement learning," *Transportation Research Part C*, vol. 97, pp. 348–368, 2018.