UDC: 621.397.63 https://doi.org/10.15407/jai2023.01.081

MODERN VIZUALIZATION TECHNOLOGIES FUSION – THE WAY TO ARTIFICIAL INTELLECTUAL SYSTEMS

L. Malinovskii
V. M. Glushkov Instituteof Cybernetics of the National Academy of Sciences of Ukraine, Ukraine
40, Academician Glushkov ave., Kyiv, 03187
LevMal@gmail.com
https://orcid.org/0000-0002-0109-928X

Abstract. To create artificial intelligent systems that receive information in the form of images: industrial robots, autonomous vehicles, modeling of objects or the environment, video surveillance, video data can be presented as a sequence of images from various cameras or sensors, of which lidars and cameras are the most studied and discussed. The work reveals and analyzes the possibilities, advantages and disadvantages of these computer vision technologies. Examples are given and researched ways, directions and prospects for the development and improvement of lidar systems. It is shown what new possibilities for autonomous intelligent unmanned systems are opened by the combined usage of cameras and lidars. Such fusion makes it possible to use the advantages of both technologies, to solve problems that seemed insoluble yesterday. It is clear that the symbiosis of these two devices, which work in real time, is crucial for many applications such as autonomous driving, industrial automation and robotics. Especially in the case of autonomous vehicles, efficient fusion of data from these two types of sensors is important for object depth detection as well as object recognition at short and long distances. Since both sensors are capable of simultaneously capturing different environmental attributes, integrating these attributes with an efficient data fusion approach greatly improves reliable and consistent environmental perception.

Keywords: lidar, videocamera, autonomous vehicles, virchual and augmented reality.

Computer eyesight or computer vision — theory and technology of creating machines that can detect, track and identify objects.

As a scientific discipline, computer vision belongs to the theory and technology of creating artificial systems that receive information in the form of images. Video data can be represented in many forms, such as a video sequence, images from different cameras or sensors. As a technological discipline, computer vision seeks to apply computer vision theories and models to the creation of computer vision systems. Examples of such systems can be:

- process management systems (industrial works, autonomous vehicles);
- video surveillance systems;
- information organization systems (for example, for indexing image databases);
- object or environment modeling systems (medical image analysis, topographic modeling);
- interaction systems (for example, input devices for human-machine interaction systems).

Computer vision can also be described as a complement to biological vision. In biology, the visual perception of humans and various animals is studied, as a result, models of the work of such systems in terms of physiological processes are created. Computer vision, on the other hand, studies and describes computer vision systems that are implemented in hardware or software. The interdisciplinary exchange between biological and computer vision turned out to be quite productive for both scientific fields. [1]

One of the new fields of application of computer vision, which is actively developing, is autonomous vehicles. The level of autonomy is measured from fully autonomous (unmanned) to vehicles where systems based on computer vision support the driver or pilot in a variety of situations. Fully autonomous vehicles use computer vision for navigation, that is, to obtain information about their position, to create a map of the surrounding environment, to identify obstacles. Some car manufacturers are demonstrating autonomous driving systems, but the technology has not yet reached the level where it can be put into mass production.

There are many important sensors for self-driving cars, but the most researched and discussed are lidars and cameras.

Lidar is a technology that uses time (flying principle of distance determination), known since the 1970s. In lidar systems, the sensor creates laser pulses. The distance to the object is then calculated based on the time it takes for the beam to travel back. Lidar offers very high resolution with long range and wide field of view. As a result, the laser sensor is able to detect even non-metallic objects at long distances, such as stones on the road, with a high degree of reliability.

The advantage of the lidar over other sensors is that the device determines a distance of up to two centimeters. For comparison, GPS error is two meters. Moreover, some materials do not reflect radio waves (such as rubber), so radar, unlike lidar, will not detect a tire on the road. Also, the operation of the device, unlike cameras and radar, does not deteriorate in low light and bright light conditions. He always sees the surroundings equally well. An infrared diode or laser is used as an active source, and a light-sensitive receiver is located nearby.

The established translation of LIDAR as "laser radar" is not entirely correct, because in short-range systems (for example, intended for indoor operation), the main properties of a laser: coherence, high density, and instantaneous radiation power are not in demand, light emitters in such systems can use ordinary LEDs. However, in the main fields of application of the technology (atmospheric research, geodesy and cartography) with radii of action from hundreds of meters to hundreds of kilometers, the use of lasers is inevitable. [2]

The advantages of lidar include: extreme reliability, very close to 100%, when detecting various objects around and calculating their size, position, distance, and, if necessary, speed, ease of selecting objects in the sensor's field of view. Lidar uses emitted light, so it works independently of ambient light. Day or night, cloudy or sunny, cloudy or sunny - the lidar sees almost the same in all conditions. It is resistant to interference and has a higher resolution than radar.

However, it should be noted that this technology is not without some drawbacks: initially lidars were very expensive. High-resolution lidars were produced in small quantities and cost more than cars (newer models are appearing for less than $1,000).

Quite modest resolution. The best devices receive an image of 128 pixels in a vertical scan with a frequency of 10 Hz.

The range is limited. Average lidars can see up to 70,100 meters and receive less feedback from large objects such as cars at a distance of about a hundred meters. Some apply for work up to 200 meters. 1.5 micron lidars, which are even more expensive, can see further.

Most lidars had moving parts to scan the world. Flash lidars do without moving parts, but nowadays they are even more expensive (in solid-state lidars of the new generation, the number of moving parts is reduced to a minimum, or they are completely eliminated).

The refresh rate is usually lower. Moreover, as the leader scans the scene, it is distorted by the movement of the scanned cars and other objects, and because different edges of the scene are scanned at different times, a shift occurs.

Lidars can experience problems in heavy rain, snow and fog, although other light-based sensors, including cameras, behave similarly. Lidars can also sometimes be triggered by subtle things like exhaust fumes. In cars, it is better to mount lidars externally. They need every photon, so don't mount them behind the windshield.

Camera-based systems behave like humans. One or more cameras watch the scene, and the software tries to do the same thing as a human – imagine and understand a three-dimensional world from a two-dimensional image. Humans are able to transform observed two-dimensional images into a three-dimensional model of the world, and do so much better after examining the scene and observing motion parallax. Computers are currently modest in the analysis of static images and only occasionally resort to the use of motion.

People use stereo vision, but can also drive with one eye closed or missing.

Cameras are really inexpensive. Equipment costs only tens of dollars, it can be quite a lot. Because cameras are sensitive to visible light, they can see any distance during the day, as long as they have a narrow enough field of view. At night, they must use additional light. Lidars perceive shades of gray in the infrared range. Cameras see colors.

If cameras are not steerable, they have no moving parts; otherwise, they can obtain high-resolution images even for distant objects. Even widely available, there are cameras with very high resolution - while the lidar sees 64 lines, the camera sees 3,000.

Because of their high resolution and color, the cameras are able to make inferences about scenes that cannot be obtained from a low-resolution lidar image.

Cameras can see traffic lights, dimensions, turn signals and other light sources. Cameras are great for reading characters.

However, cameras also have some drawbacks. Today, computer vision does not work well enough to find all the important characteristics with the reliability required for safe driving, for example. Cameras must work with changing lighting. Objects under observation are often subject to moving shadows, and may also be illuminated from any direction (or not illuminated at all). At night, cameras need additional lighting, and headlights may not be enough. Cameras are affected by weather and other environmental conditions, so they cannot be relied on all the time. This limitation was discovered in an accident where a car with a partial set of autonomy tools collided with a white truck that was not recognized by the camera against a background of white clouds. Computer vision tasks require high-performance processors or specific chips to work at the level of current requirements. Given that the use of cameras requires algorithmic breakthroughs, it is difficult to predict when they will be good enough for self-driving. In turn, the lidar can create a complete 3D map of the scene in a single pass. Multiple passes can improve the picture and help him, including to estimate the speed.

A big advantage of lidar is also that at least in the case of objects of a decent size, such as pedestrians, cars, cyclists and large animals, the laser beam will always be returned, indicating their presence. The system may not be able to figure out exactly what it is, but it will know that the object is there and will be more and more confident as it gets closer. If something large is blocking the road in front of you, you must stop no matter what it is, although there are exceptions - birds or debris blown up by the wind. Within a certain range of distances and sizes, lidar is close to 100 percent accurate, and that's very important.

The synergy of camera and lidar allows, by combining the advantages of both technologies, to solve problems that seemed unsolvable yesterday. It is clear that the symbiosis of these two real-time devices is crucial for many applications such as autonomous driving, industrial automation and robotics. Especially in the case of autonomous vehicles, efficient fusion of data from these two types of sensors is important for object depth detection as well as object recognition at short and long distances. Since both sensors are capable of simultaneously capturing different environmental attributes, integrating these attributes with an efficient data fusion approach greatly improves reliable and consistent environmental perception.

The pioneers of virtual and augmented reality (VR&AR) technology believe that the best solution for capturing real content for virtual reality will be lidar-based spatial scanning technology. The technology is based on the interaction of laser and photography, forming a complete 3D model of space. The construction of the model, without delving into technical details, is based on 3 stages:

– Laser cameras perform spatial analysis with high accuracy, determine the distance (depth) of each element in space from the position of the cameras.

– On the basis of the received data of in-depth analysis, the construction of a 3D model, "casts of space" is made.

– "Mold" acts as a base for superimposing a photograph.

As a result, the source content presents a three-dimensional form and makes it possible to actually move inside the scanned space. [3]

The renaissance of lidar technology began when his concept helped teams win the DARPA Urban Challenge in 2007. Since then, lidar systems have practically become the standard for robocars. By 2021, according to the IDTechEx review (idtechex.com), 106 manufacturers produce 156 lidar products, and this market segment is actively developing due to investments in independent forward-looking research in the field of the latest technologies. The growth of the lidar market is almost six times higher than the growth index of the world economy and, according to forecasts, should grow three times to $3 billion in 2025.

You should pay attention to three essential factors that distinguish lidars from different manufacturers: how the laser is directed in different directions, how the time for the journey there and back is measured, and the light of which frequency is used.

Most leading lidar manufacturers use one of four methods of directing laser beams in different directions.

Rotating lidar. The advantages of this approach are 360-degree coverage, but critics question whether it is possible to make a cheap and reliable rotating lidar suitable for the mass market.

Mechanical scanning lidar uses a mirror to redirect a single laser beam in different directions. Some companies are using an approach called a "microelectromechanical system" (MEMS) to control the micromirrors.

The antenna array is active in phase with a victorious series of viprominuvachiv, building changes directly to the laser menu, changing the phase of the signal between the sudanim transmissions.

The flash-based lidar illuminates the entire area at once. Existing technologies use a single wide-angle laser. The technology struggles over long distances because only a small fraction of the laser light reaches any given point.

Distance measurement: lidar measures the time it takes for light to reach and bounce off an object.

There are three simple ways to do this: Time on the road. The lidar sends out a short pulse and measures how long it takes for the return pulse to be fixed.

Continuous emission lidar with frequency modulation (CFM). Sends a continuous beam of light, the frequency of which constantly changes over time. The beam splits into two, and one of them goes to the outer world, and then, after returning, unites with the other. Since the frequency at the source of the beam changes continuously, the difference in the path of the two beams is expressed in terms of the difference in their frequencies. The result is an interference pattern whose reflection frequency is a function of travel time (and therefore distance). This path may seem complicated, but it has several advantages. The CFM lidar is resistant to interference from other lidars or the sun. The CFM lidar can use Doppler shift to measure the speed of objects, not just the distance to them.

Continuous emission lidar with amplitude modulation (CAM) can be considered as a compromise between the two previous options. Like a simple travel-time sensor, lidar sends out a signal and then measures the time it takes to get there and back. But where simple systems send out a single pulse, the CAM lidar sends out a pseudo-random stream of digital zeros and ones. Proponents of the approach say this makes the CAM lidar more resistant to interference.

Let's move on to the wavelength of the laser. Lidars from well-known manufacturers use one of three wavelength options: 850, 905 or 1550 nm. This choice is important for two reasons. One of them is eye safety. The fluid inside the eye is transparent to light with wavelengths of 850 and 905 nm, which allows light to reach the retina. If the laser is too powerful, it can cause irreparable damage to the eye. On the other hand, the eye is opaque to radiation with a wavelength of 1550 nm, which allows such lidars to operate at higher power without harming the retina. Increasing the power allows you to increase the range of action. So why isn't everyone using 1550nm lasers in lidars? Detectors operating at frequencies of 850 and 905 nm can be created on the basis of inexpensive and common silicon technologies. Creating a lidar with a wavelength of 1550 nm requires the use of exotic and expensive materials such as gallium indium arsenide. And while 1550 nm lasers can operate at higher power levels without posing a threat to the eyes, these power levels result in reduced range and reduced energy efficiency of the machine.[4]

The video data stream from the camera and information from the lidar (in the form of a "cloud of points") combined at the "iron" level will allow solving previously intractable tasks in the following areas:

- virtual and augmented reality;
- autonomous vehicles;
- monitoring of climatic changes;
- archeology;
- geodesy.

Such a synthesis requires finding the intersection of the lidar field of view and the image from the camera, and then assigning the values of the "cloud of points" that coincide with the pixels of the image for further calculation of the depth of all pixels of the scene.

In the fall of 2020, a landmark event took place almost imperceptibly: the presentation of the iPhone 12 Pro and iPhone 12 Pro Max, which just received a lidar sensor on the rear camera unit. In this way, a 3D dimension is added to the everyday practice of interacting with a computer.

The first 3D lidar was introduced by Velodyne more than ten years ago. The rotating device cost about $75,000 and was much larger than a smartphone. Apple needed to make the lidar cheap and small enough to fit into the iPhone, and vertical-beam lasers allowed the company to achieve that. The lidar in the smartphone sends light using an array of Vertically Emitting Lasers (VCSELs) manufactured by Lumentum. It then captures the return flash using an array of single-photon avalanche diodes (SPADs) supplied by Sony. These technologies (VCSEL and SPAD) are being used by Quster, Ibeo, Sense Photonics and others to create a much more powerful lidar for the automotive market. Vertically emitting lasers and single-photon avalanche diodes are interesting because they can be mass-produced using conventional semiconductor device manufacturing technologies. So, the benefit comes from huge savings on high-volume production. As sensors based on vertically emitting lasers become more common, their quality (and price) will increase.

Because VCSELs emit perpendicular to the substrate surface, many lasers can be placed on a semiconductor die. The technology has been around for a long time, but it has always been considered not powerful enough for use in lidar. But Ouster says it knows how to make a high-efficiency lidar using VCSELs and has announced plans to release a new solid-state lidar with no moving parts. Instead of lining up anywhere from 16 to 128 lasers, Ouster's new device will use 20,000 vertically-emitting lasers arranged in a two-dimensional grid. In addition, Ouster uses another semiconductor technology mentioned above, single-photon cascade diodes (SPADs), to detect the returned light. As the name suggests, they are sensitive enough to find a single photon. High sensitivity means they suffer from noise. In order to use such diodes in devices such as lidars, complex post-processing is required. Like VCSELs, SPADs can be fabricated using standard silicon chip manufacturing techniques, and many SPADs can be placed on a single die. This made it fairly easy for Ouster to go from 64 laser devices last year to 128 lasers, which were announced in January and will begin shipping in the summer. The companies simply replaced the chips with 64 lasers and 64 detectors in the old model with new 128 chips. In the coming years, Ouster, Ibeo and Sense must decide how to develop the performance of the combination of vertical-beam lasers and single-photon avalanche diodes so that their devices can operate with a range of 200 meters. If they succeed in solving this challenge, the low cost and simplicity of the chips will give these companies a decisive advantage in the automotive industry.[5]

Companies that do not use vertical lasers are also moving into this market. One of the most prominent companies in this field is Luminar, which announced a partnership with Volvo in May. Volvo plans to release cars with lidar from Luminar in 2022.

All these designs have their strengths and weaknesses (and they are different). So far, Luminar can boast of a significant range - as much as 250 meters. It is possible that Luminar uses lasers with a wavelength of 1550 nm, which is far outside the range of visible light. The fluid in the human eye is impervious to such light, so Luminar can use powerful lasers that will be safe for the human eye. Luminar lidars also have a wider field of view than Ouster devices. The biggest question for Luminar is whether they will be able to meet the $1,000 price tag.

Lidar from the innovative company AEye has a lot in common with Luminar. It uses a mechanical scanning mirror to direct the eye-safe 1550nm laser beam, allowing it to operate at higher energy levels. As a result, the lidar from AEye has impressive range characteristics. AEye says their lidar can see up to 1,000m away – far more than the 200-300m range boasted by the most expensive devices. Most lidars use a fixed scanning pattern. AEye's lidar uses a different approach, which the company calls moving scanning. The AEye scan pattern can be configured programmatically and changed dynamically. The dynamic scanning scheme works with the flexibility of a fiber laser. The software controls not only when the next measurement will take place, but also how much energy will be used and therefore how far the next measurement will be made. As a result, when the lidar detects an object that is far away, it can increase the scanning resolution and energy level in that part of the image, and acquire more data points. The result can be a high-resolution scan that can help distinguish a pedestrian, a motorcycle, or bulky debris left on the road.[6]

Such a combined sensor should outperform both cameras and lidars separately. AEye startup has raised $61 million in investments to implement its plan.

According to some forecasts, VentureBeat writes, by 2030 there will be 10 million driverless cars on the roads. Most of them will be able to navigate in space only thanks to lidars — devices that, emitting laser waves, determine distances to surrounding objects.

The lidar market is growing along with confidence in an unmanned future. It is expected that in five years it will already reach $3 billion. But although lidars will be installed on many cars, they all differ in the quality and detail of the digital map. High-quality devices usually cost several thousand dollars, but they still have disadvantages. For example, they are weak in determining the type of objects.

However, less accurate in determining the range of smart video cameras perfectly distinguish between images. AEye combines these two technologies, claiming that for the first time, the video stream and lidar information are combined at the "iron" level.

The RGB channel of the camera is combined with the lidar. As a result, a zone with a radius of 300 meters is formed, in which the lidar and camera hybrid recognizes all objects surrounding the robot.

But the advantages do not end there. Beyond the 300-meter zone, the lidar from AEye works in standard mode, but "hits" at 1000 meters - 4 times farther than standard devices. At the same time, the developers say that they are potentially able to increase the range to 5-10 km.

The above results show that the synthesis of lidar and camera data goes far beyond the simple sum of their capabilities, and further innovations should be expected in the future integration of these science-intensive technologies.

References

1. https://en.wikipedia.org/wiki/Computer_vision/
2. https://habr.com/ru/company/itelma/blog/479736/
3. https://www.geospatialworld.net/blog/
4. https://www.mdpi.com/2073-8994/12/2/324/htm
5. https://www.sciencedirect.com/science/article/pii/S1877050921005767
6. https://habr.com/ru/post/441706/

The article has been sent to the editorial office 29.09.22.

After processing 07.11.22.