Unmanned Vehicle LIDAR-camera Sensor Fusion Technology

Kuzin V.D., Pronin P.I., Zarubin D.N., Shashkina K.M.

Nizhny Novgorod State Technical University n.a. R.E. Alekseev, 24, Minina St., Nizhny Novgorod, 603950, Russia
E-mail: chromium32@mail.ru

Abstract
With the growth of computing power and the development of information technology in the last few years, the tendency for the development of unmanned vehicles has increased. Many leading corporations sponsor this industry. Since the creation of the concept of an unmanned vehicle implies that the computing systems in its composition must make decisions about the method and direction of its further movement, predict the further actions of the traffic participants and move in accordance with the rules of the road, these vehicles are equipped with powerful telemetry systems that allow to accurately evaluate surroundings. But, since there is a problem in incomplete data from one device (sensor), in this article, using a specific example, sensor fusion technology is considered that allows combining data from several sensors to achieve maximum results in solving problems of developing an unmanned vehicle.

Keywords: LIDAR, camera, sensor fusion, unmanned vehicle

1. Unmanned vehicle
If we consider the telemetry system of an unmanned vehicle, then it can be divided into sections corresponding to a specific task in solving the problem of unmanned traffic. For example, the localization and positioning system and the inertial system make it possible to understand the position in space of an unmanned vehicle.

Groups of optical sensors and radars allow to learn about the parameters of obstacles in the path of movement of an unmanned vehicle and collect data about the surroundings and other participants in the movement. An important factor in the implementation of an unmanned vehicle system is the ability to identify obstacles along the path of an unmanned vehicle using certain sensors or a group of sensors and take actions to prevent a collision. Most systems use obstacle detection algorithms using modular stereo cameras, deep vision systems. In this configuration, an active laser rangefinder of the optical range - LIDAR - is used as an obstacle detection sensor.

The principle of operation is similar to the principle of the radar - a directional signal, and in this case a light laser beam, is emitted by the device, reflected from the object and captured by a highly sensitive sensor. The response time of the signal is directly proportional to the distance to the object in the field of view of the device. Light waves are subject to scattering in any medium, including in the air, therefore it is possible not only to determine the distance to dence (light-reflecting) discrete targets, but also to fix the light scattering intensity in transparent media, which makes it possible to determine in some way such parameters of an object as its illumination, texture, in some cases even color. This article discusses the application of sensor fusion technology to a group of optical sensors, namely, the LIDAR bunch and camera used in the telemetry system of an unmanned vehicle based on the Gazel “Next” (figure 1(a)) and Gazel “Next-drive” (figure 1(b)).
1.2. Optical Sensor Group
In this configuration, LIDARs from Velodyne are used. Testing and implementation of the technology was carried out using VLP-16 models[1]. The standard ROS[2] driver module for working with LIDAR Velodyne VLP-16 (figure 2) publishes data from the device by topic in the ROS framework used in development.

Figure 2. Velodyne VLP-16

The input data is the so-called point cloud, which is a three-dimensional map of the sensor’s environment in its field of view. In the current implementation and production of LIDAR, a three-dimensional map with a horizontal viewing angle of 360 ° is built on the hull of an unmanned vehicle. Each point of the cloud contains the necessary and sufficient information for further calculations, namely:

a. Intensity. Intensity - the degree of reflection of the sensor beam from the surface, for example, from a solid light surface, the reflection intensity will be higher than from a black surface.
b. Coordinates in space. Each point has a spatial coordinate in the format (x, y, z) relative to the sensor itself.
c. Belonging to the ring. VLP-16 is a sixteen beam sensor; accordingly, 16 light (laser) emitters rotating in a circle are provided at the physical level of the sensor design.

The Basler[3] acA1300-200uc camera (figure 3) was also used. This camera, equipped with the necessary lens, allows you to receive high-quality video images with a high frame rate, which is important when installing on an unmanned vehicle.
2. Preparatory stages

The sensor fusion task itself involves combining data from multiple sensors. But in the process, it is important to take into account the peculiarity of the operation of these sensors and carry out the necessary preparatory steps. When working with the camera and LIDAR, one of these steps is calibration.

2.1. Camera calibration

Before making measurements in the field of computer vision, it is worth considering the distortion of the image by the camera itself. Calibration (figure 4) corrects complex distortions, which are determined by the shape of the lens, errors in its mechanical fixation, errors in the manufacture of the sensor. Further steps to combine data and image processing should be carried out only after calibration, otherwise calculation errors and associated distortions are inevitable. Calibration of the camera is carried out in order to find out the internal parameters of the camera, as well as the projection matrix used in the further algorithm for joint calibration of LIDAR and the camera.
After successfully performing camera calibration, having a calibration matrix, you can get a corrected image without the distortion effect.

2.2. Calibration of LIDAR and camera

Calibration of the camera and LIDAR consists in finding the correspondence of their fields of view and the correct setting. In contrast to the known methods of automatic calibration using a perforated panel[4], it was decided to develop another solution to this problem for a more optimal approach. This solution is the development of manual calibration of LIDAR and the camera. This software uses the PCL[5] library and the LIDAR point cloud processing functions it provides for the subsequent transformation of point clouds. Selecting the transformation parameters (yaw, pitch, roll) of the spatial vector of the input cloud and the displacement along the coordinate axes, we can accurately correlate the data of the point cloud on the spatial position of the object in the LIDAR field of view with the position of the object on the camera frame.

3. Sensor fusion

When the sensors are calibrated, the data merge can be performed directly. The combined fields of view of the two sensors allow to expand the range of detection of objects. So, for example, specifically in this case, SF is used to determine the physical parameters of objects in the field of view of these devices with their detection and recognition by the neural network in the image.

Also, the used functionality allows you to directly visualize LIDAR data on the image frame in accordance with the intensity at the current point (figure 5(b)). This approach helps with preparatory steps when it is necessary to find the correspondence of the field of view of two sensors.

How it works. The image from the streaming video from the camera is fed to the input of the neural network. That in turn determines the objects and passes the list of detected objects further to the SF module along with the coordinates of the boundaries of the detected objects on the frame. Further, the module, having received data on the boundaries of objects in a 2D image, draws a correspondence between the object in the 2D field of view of the camera and the 3D view of LIDAR. After completing these steps, the output gives the physical parameters of the desired object directly in 3D - the distance of the object, its size, etc. This approach allows you to work with segments of the LIDAR point cloud as separate objects of space and classify them according to the classes that the neural network provides at the recognition stage. As one of the stages of this work, we can identify the limiting volumes (figure 5(c)) of objects immediately in 3D and use the identified boundaries along the unmanned vehicle’s path, for example, to correct the trajectory of its further movement.
Figure 5. Steps of sensor fusion:
(a) clear frame, (b) projected neural network data and LIDAR data to the frame, (c) visualized at Rviz limiting volumes of objects

Future work
The immediate prospects are optimization of matching algorithms and matching projections of sensor data, increasing processing speed and improving visualization algorithms. The described sensor fusion methods and algorithms are implemented in the unmanned vehicle control program complex and are currently being tested.

This research done with the financial support from Ministry of Science and Higher Education of the Russian Federation (the unique project identifier is RFMEFI57717X0268).

References
[1] Velodyne VLP-16. Website (https://velodynelidar.com)
[2] ROS – Robotic Operation System. Website (https://www.ros.org)
[3] Basler camera product. Website (https://www.baslerweb.com/en/)
[4] but_velodyne ROS package. Website (http://wiki.ros.org/but_velodyne)
[5] PCL – pointcloud library. Website (http://pointclouds.org)