XLIII CILAMCE
21-25, November, 2022
FOZ DO IGUAÇU - PR - BRAZIL

Measurement of classified points using stereo vision and techniques of segmentation in disparity map for detection of obstacles.

L. Pin[1], R. Vivacqua[1], M. Cuadros[1]

[1]*IFES - Federal Institute of Espírito Santo*
Serra, 29166-630, ES, Brazil
luizabrpin@gmail.com, rafsat@ifes.edu.br, madsouzacuadros@gmail.com

Abstract. This work presents an obstacle detection system using a stereo camera as a sensor and segmentation techniques in occupancy maps identifying navigable and non-navigable regions. The disparity images are used to build a 3D point cloud whose distances are calculated from the ground plane. Then, points that are in a range of distances are classified as potential obstacles and recorded on an occupancy map. Clustering techniques are then applied to identify obstacles to identify the position, size and speed of obstacles. To validate the algorithm developed in this work, some real tests were carried out with a full-scale autonomous vehicle.

Keywords: self-driving car, disparity map, obstacle avoidance.

1 Introduction

According to the Ministry of Transport of Brazil, the numbers of accidents on urban roads are increasing every year. Recklessness, road conditions, alcoholic beverages, adverse weather conditions, inattention and drowsiness while driving cause dramatic loss of life and annual losses worldwide [1].

According to the Brazilian Institute of Geography and Statistics - IBGE, between 2012 and 2017, the number of elderly people grew in all units of the federation, the population aged 60 years and over was 25.4 million [2], they have There is a world population that is aging, limiting its autonomy due to accessibility difficulties and insecurity behind the wheel. The effective application of these technologies can contribute in a decisive way to increase safety on public roads, drastically reducing the accident rate and bringing safety in vehicle traffic.

Today, car manufacturers offer many advanced driver assistance systems in new vehicles they market. At the same time, they are actively working on Advanced Driver Assistance Systems (ADAS) applications as well as autonomous driving capabilities. It has also become increasingly important to provide adequate resources for pedestrian detection and animal accident prevention. These developments not only improve the safety of pedestrians, but also of vehicle occupants [3].

Technologies such as Park Assist, which automatically parks the car in parallel or perpendicular spaces, making the goal much more practical and safer are becoming common in popular cars. This automatic parking technology uses sensors around the car to detect the size of the space and calculate the need to maneuver. Another example is the SPARC - Secure Propulsion using Advanced Redundant Control project developed by a consortium of automotive companies and European universities, with the aim of improving traffic safety and efficiency for heavy goods vehicles and passenger vehicles, this system uses the SDCS System - Safety Decision Control, which is designed to check the probability of driver failure and react accordingly [4].

Therefore, it is clear that the evolution of mobile robotics technology has brought a great contribution to increase road transport safety with the help of ADAS. It is also important to emphasize that the consolidation of ADAS technology serves as the basis for the next stage of evolution of transport systems, that of autonomous vehicles. This work presents an obstacle detection system using a stereo camera as a sensor and segmentation techniques in occupancy maps identifying navigable and non-navigable regions. The effective application of this experiment can contribute to increase safety on public roads, drastically reducing the accident rate.

2 Theoretical Foundation

Obstacle detection by stereo vision is a set of computational and mathematical methods capable of extracting information from the environment and using it to increase the driver's level of safety and comfort. In the field of autonomous vehicles and ADAS systems, operating conditions require sensors to have a range in a very wide range, from a few meters to tens of meters. The types of sensors capable of operating in this condition are cameras, radar and LIDAR. All are based on the emission and/or reception of acoustic and/or electromagnetic waves.

Radar is the longest range sensor, being the only one that does not go blind in foggy conditions. Its principle and operation is the same as the radars used in aviation, that of emitting a radio signal and detecting its reflection. The big issue with Radar is that it only detects metallic bodies (other cars), and as it only emits a beam, it doesn't allow a detailed perception of the environment. For this reason, its main application is in adaptive cruise control.

LIDAR is a sensor used in mobile systems (autonomous or non-autonomous) and a direct method of data capture. The sensor emits laser beams in the infrared band and is able to model the terrain surface three-dimensionally. This technique is used for topographic surveys, to characterize the vegetation structure, as well as the volumetry of buildings and urban environments more quickly and reliably. Object detection with the acquisition of depth information for each detected object is an important issue for robotic or vehicular navigation, providing information about possible obstacles and their location in relation to the vehicle or robot. LIDAR is one of the most widespread of these sensors to request this information, through a laser that is used for remote sensing and using the reflected light to measure distances between the sensor and target objects, such as obstacles, other vehicles and road irregularities [5].

Cameras have the advantage of being inexpensive, generating detailed information, having a good field of view and measuring over a wide range of distances. They can operate individually (monocular vision) or in pairs (stereo vision). In monocular vision, obstacles are normally detected through pattern recognition, that is, in the interpretation of the shapes and regions contained in the image. This type of work requires a lot of computational resources and the best results are currently obtained with deep neural networks. In stereo vision, obstacle detection is done by measuring the distance of thousands of points in the captured image, using the triangulation technique, which is a technique that requires less processing than pattern recognition by deep neural networks.

2.1 Stereo View

Stereo vision is composed of two capture sources, cameras and lenses, parallel or not, and spaced, which acquire two-dimensional images, containing slightly different contents due to the positioning of the cameras and the incidence of light [6]. The stereo camera simulates the human vision system, the two horizontally positioned lenses are offset and capture two similar images. Stereo cameras have some advantages over LIDAR: much lower cost, greater mechanical robustness, can be mounted internally, ability to detect colors and textures. All these features make stereo vision the most used sensor in ADAS systems currently on the market. The main disadvantages of stereo cameras over LIDAR are the small field of view and the inability to measure distance in non-textured regions, although this condition is very rare in the real world.

Epipolar geometry is the geometry that describes stereo vision, it establishes the mathematical relationship between two points projected on two displaced image planes. Once the stereo camera pair is calibrated, that is, the rotation and translation between them are known, it is then possible to perform a 3D reconstruction of a scene from a pair of corresponding points. The side-by-side arrangement is the most commonly used in stereo cameras. It has the great advantage of generating horizontal epilines, which speeds up the point matching process. A perfectly aligned side-by-side mount (parallel image planes) is something very difficult to achieve in mechanical camera attachments. What is then done is a process of geometric transformation of the images, called stereo rectification, so that this condition is perfectly achieved by software. Fig. 1 a) illustrates the grinding process. The gray planes correspond to the real planes of the cameras which are not perfectly parallel. In yellow we have the plans rectified by software (homography).

While depth recovery can be done with any pose difference between cameras, the rectification condition makes the stereo matching process much easier as the search is restricted to a horizontal line. The horizontal displacement (in x) corresponding to a given piece of image seen by the camera on the right in relation to the one on the left is called disparity and is inversely related to the distance of the object from the cameras plane.

The pair of images, coming from side-by-side cameras, presents a displacement between relative positions of local parts depending on the distance that these local components are from the sensor, knowing the difference in the position of a point between one image and another it is possible to measure depth, similar to the way human vision perceives depth [7].

Disparity is the horizontal difference between corresponding points in different images, measured in pixels. This disparity effect can be noticed when we observe objects very close to our eyes and we alternate between them, closing one and opening the other and vice versa. This "displacement" effect is the disparity, and the further away the object is, the smaller this effect.

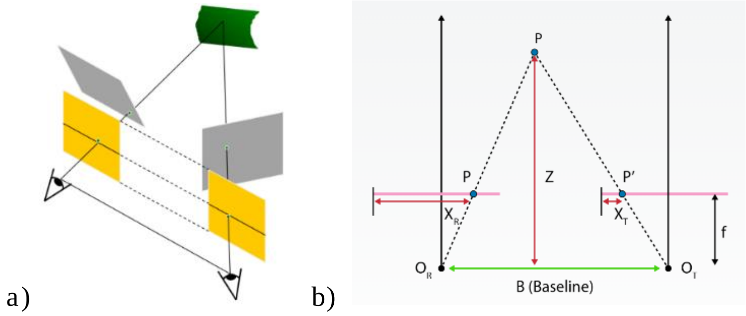


Figure 1. a) Epipolar geometry ; b) Disparity calculation.

The extraction of the depth of a point in relation to a stereo camera is based on the triangulation between two parallel image sensors whose baseline (B) and focal plane (f) are known. The distance illustrated in Fig. 1 b) is obtained from the relationship of triangles, and can be expressed by the following equation:

$$Z = \frac{f * B}{X_l - X_r} \qquad (1)$$

The disparity map is an image where the intensity value of each pixel corresponds to the disparity value. The depth match, or disparity value, is the value of the horizontal coordinate difference between two corresponding pixels of the two images of the vertical and horizontal component of the disparity matrix.

2.2 ADVANCED DRIVER ASSISTANCE - ADAS

According to ADASE - Advanced Driver Assistance Systems in Europe what is called ADAS - Advanced Driver Assistance is a set of systems and subsystems on the way to a fully automated road system. The objective of this system is that the driver error is reduced or eliminated and its efficiency is improved, since its functionality is to facilitate the performance of the drivers' tasks, providing advice, instructions and warnings in real time [8]. The main functions of ADAS are: Forward Collision Alarm, Lane Departure Alarm, Speed Sign Recognition, Pedestrian Detection and Traffic Jam Assist.

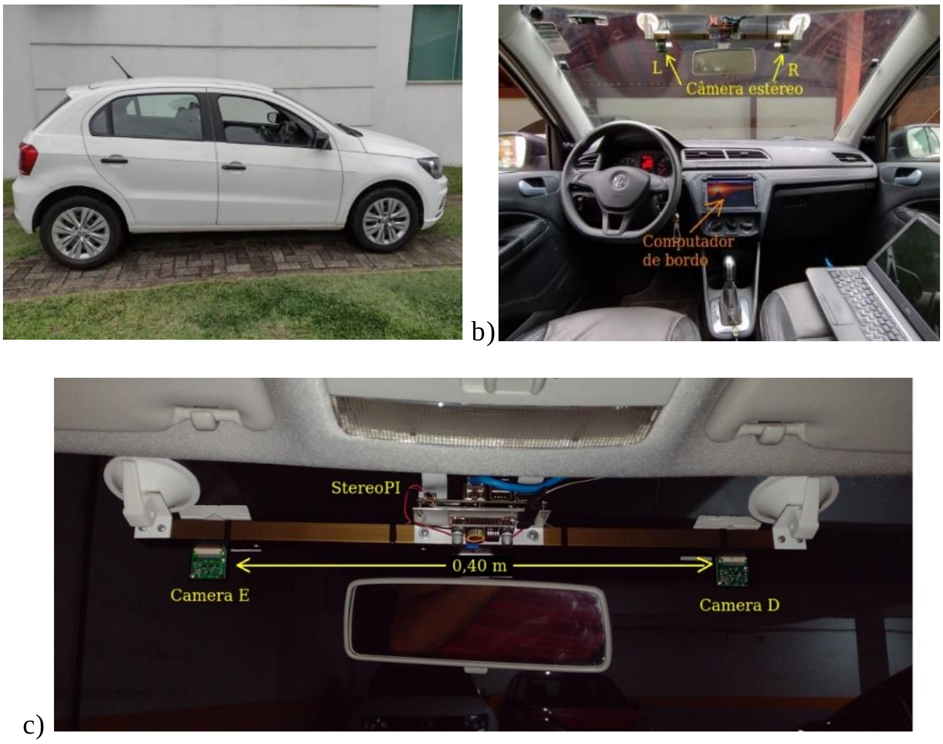
ADAS was developed for car users to avoid unexpected accidents. This system is equipped with many features to support the safety of passengers, drivers and surroundings by introducing a feature called pedestrian detection. Pedestrian detection systems use sensors such as LIDAR, stereo vision camera and radar, and can have data fusion from different sensors. In a camera-based pedestrian detection system, it can get rich texture information provided that can be used not only for pedestrian detection, but also for analyzing pedestrian posture and intent in the advanced pedestrian safety function. The camera is the sensor most commonly used as the basis for a pedestrian detection system [9].

3 Methodology

3.1 Test vehicle feature

To carry out the validation tests of the system proposed in this work, a real car was used, a Volkswagen GOL 1.6, properly equipped with computers, sensors and actuators. The hardware architecture was developed during

[10] in the area of autonomous vehicles. In Fig. 2 (b) there is an overview of the interior of the vehicle, where a stereo camera and the on-board computer monitor can be seen. The notebook in the photo was used simply to remotely access the on-board computer.



a) b) c)

Figure 2. Test vehicle: (a) Photo of vehicle. (b) Internal view. (c)

The camera on the test vehicle is a StereoPi v1 Fig. 3 (b), a Raspberry Pi-based stereoscopic camera with an open source. Because it is sold in kit form, it has a very attractive price without sacrificing quality. A great advantage of being marketed as a kit is that it allows the user to mount the camera in the configuration they prefer, vertical (top and bottom) or horizontal (side by side). Also, the base distance can be chosen according to your application. StereoPi is capable of capturing, recording, streaming and processing stereoscopic video (disparity calculation) in real time.

3.2 Collecting test images

The capture of images to compute the disparity map took place on a stretch of the north south highway, in a condition of intense traffic. The road chosen for the collection of images is characterized by having an intense flow of vehicles and presenting a flat road with defined lanes painted on the ground, this condition is ideal to reproduce the operation of ADAS. The images were captured and processed by the stereo camera at 15 frames per second and the disparity maps were saved in 16-bit formats, maintaining maximum obstacle depth resolution.

3.3 Camera calibration

Camera calibration is the procedure done to determine the relative position between the camera's reference system and the ground plane, where obstacles are supported. The ground plane is calculated through analytical geometry by the three-point plane equation method. In Fig. 3 a) the selected points are shown superimposed on the image of the left camera and in Fig. 3 b) on the disparity map. The choice of the three points was made manually, selecting points on the disparity map corresponding to the tracks on the ground.

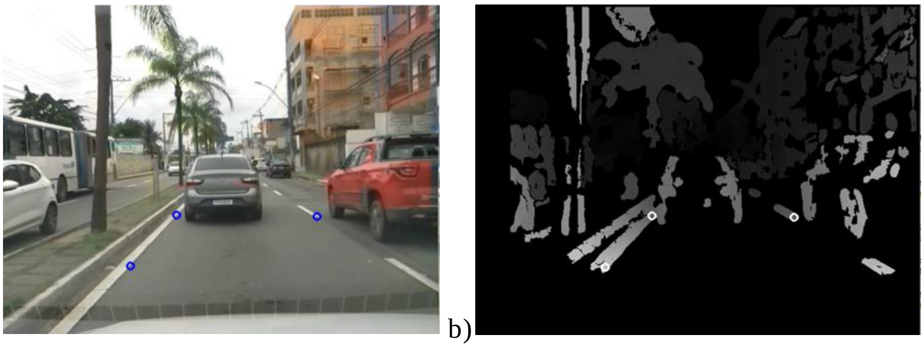
a)  b)

Figure 3. Points for calibration: (a) left camera image; (b) disparity map

4 Test and results

4.1 Obstacle detection

The first step of the obstacle detection algorithm is the selection of 3D points of interest through filtering. Filtering aims to discard points that are outside the possible range of collision, such as points higher than the car or points stuck to the asphalt. Points considered as belonging to possible obstacles are selected according to their distance from the ground plane, that is, their height. The filtering condition for the selection of points is between 0.2 and 1.6 meters, that is, for the algorithm to identify the obstacle, it is necessary that the height is within this range and that the depth component (Z) is less than 20 meters. Filtering is necessary for removing false points, differentiating small bumps and holes in the track and for optimizing the algorithm's processing.

4.2 Clustering of cores

The second step of the obstacle detection algorithm is clustering the cores. The grouping aims to connect the closest pixels forming a single obstacle. The clustering of cores, called clusters of points, was performed on the filtered image. The *connectedComponentsWithSta*s function from the OpenCV library was used to perform this grouping. The function captures groups of connected pixels in the image with 8 pixels neighborhood. The input image for the function must be binary. Fig. 4 a) shows the result of grouping the regions connected by 8 correspondents. Each cluster was identified with a unique number and placed inside a bound box.

4.3 Map of obstacles

The obstacle map is a 2D representation of the locations free or not of obstacles, in a perspective seen from the sky (bird eye view). With it, it is possible for the ADAS action planner to identify if there are obstacles in the collision region of the car (in pink in Fig. 4-b) and that need warning or braking. The red line at the base corresponds to the width of the vehicle. In the illustration of Fig. 4, the collision line is straight because it considers that the vehicle is moving in a straight line.

The obstacle map is generated by projecting the 3D points onto the ground plane. Its origin corresponds to the projection of the point (0,0,0) also on the ground plane.

The points on the obstacle map are following the grouping that was made in the image of Fig. 4 a). With this, the obstacles are no longer scattered points and start to have dimension (width and depth). This information can be useful, for example, to determine the type of obstacle (car, motorcycle, etc.).

a) b)

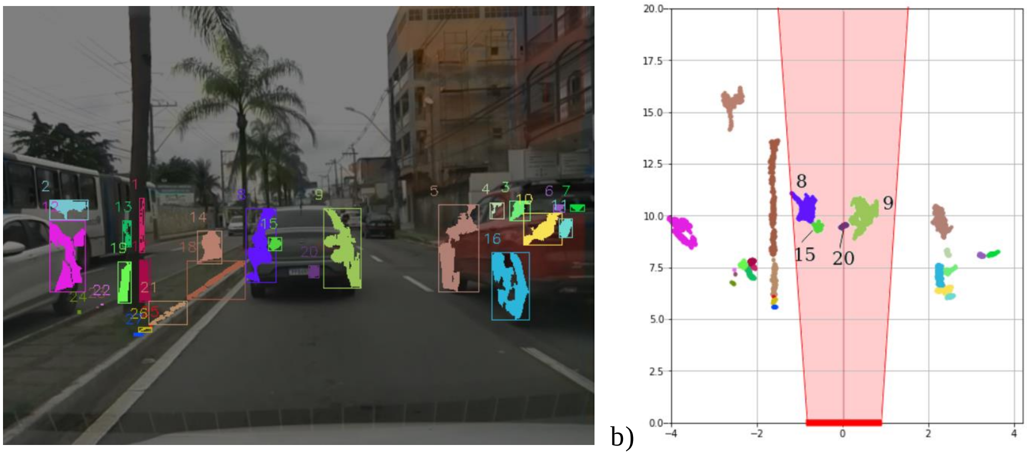


Figure 4. a) identification of obstacles by numbered bounding boxes; b) obstacle map.

5 Conclusion

This paper presents the development of an ADAS system for collision warning and automatic braking, based on stereo vision. The information captured by the camera is projected onto an obstacle map, from where the ADAS action planner extracts alert or reaction information. This document also presents the technical details of the vehicle used to capture the experimental data and the calibration procedures necessary for the correct construction of the obstacle map. In the continuation of this work, it is foreseen the tracking and estimation of the speed of obstacles, incorporation of the gyroscope in the generation of the collision region, elaboration of the alarm and reaction algorithm and experimental validation tests.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors.

References

[1] K. A. Brookhuis, D. W. Waard and W. H. Janssen, " Behavioural Impacts of Advanced Driver Assistance Systems – an overview". EJTIR, 1, no. 3, pp. 245 – 253, 2001.
[2] IBGE – BRAZILIAN INSTITUTE OF GEOGRAPHY AND STATISTICS. "Number of elderly people grows 18% in 5 years and exceeds 30 million in 2017." Available at: <http://www.ibge.gov.br/home/>, accessed on April 25, 2022.
[3] H. Estl, "Paving the way to self-driving cars with advanced driver assistance systems." *Worldwide Systems Marketing for Advanced Driver Assistance Systems (ADAS)*, Texas Instruments, 2015.
[4] SPARC - Secure Propulsion using Advanced Redundant Control. *Multi-annual roadmap for robotics in Europe*. 2015
[5] LIDAR. LIDAR - http://www.lidar-uk.com/. 2015. UK.
[6] A. Hosni, M. Bleyer and M. Gelautz. , " Secrets of adaptive support weight techniques for local stereo matching." *Computer Vision and Image Understanding*, 117:620– 632, 2013.
[7] R. Hartley and R. Zisserman; "Multiple View Geometry in Computer Vision. Second Edition.", *Cambridge University Press*, 2004.
[8] ADASE. Advanced Driver Assistance Systems in Europe. *Project reports European Commission*, 1998.
[9] F. Camara, N. Belloto, S. Cosar, D. Nathanael, M. Althoff, J. Ruenz, A. Dietrich and C. Fox. *Pedestrian Models for Autonomous Driving Part I: Low-level Models, from Sensing to Tracking*. IEEE. 2020.
[10] R. Vivacqua, "Back Lane Marking Registry: an approach to locating and following the path by autonomous vehicles on a signposted lane." UFES - Federal University of Espirito Santo, 2017.