Unmanned Ground Vehicle with Stereoscopic Vision for a Safe Autonomous Exploration

Jesús Jaime Moreno-Escobar[1*], Oswaldo Morales-Matamoros[2*], Ricardo Tejeida-Padilla[3†]
*ESIME Zacatenco, Instituto Politécnico Nacional, México.
†EST, Instituto Politécnico Nacional, México.

***Abstract*—At present there are several systems in cars that provide assistance to the driver and the tendency is that these systems are increasingly efficient and that for their operation do not require the intervention of the driver. Computer vision is relevant in this sector due to the contribution it provides, for example, through the treatment of images algorithms are designed for the detection of objects, pedestrian detection, traffic signal detection, one lane tracking and assistant parking.In this work we present an Unmanned Ground Vehicle system through Computer vision specifically Stereoscopic vision, through a sensor we obtain a disparity map that allows us to quantify the depth of each point of the captured image. The system captures an image of its environment that through internal processing of the sensor, returns a disparity map as data, which is processed by an algorithm that allows the system to navigate in a region of the space in which it is positioned for Autonomous Exploration.**

***Keywords*—Unmanned ground vehicle; stereoscopic vision; computer vision; autonomous exploration**

I. Introduction

In modern society, automobiles are of great importance and is that most citizens own a car, in 2016 5,725,574 motor vehicles were registered in circulation in Mexico City with a population of approximately 8,918,653 inhabitants. However, there are a lot of car accidents every day. These accidents in many cases are due to lack of driver attention, fatigue, drowsiness, consumption of harmful substances, stress caused by living in urban areas with overpopulation such as Mexico City, among other factors. And in 2016, 11,449 land-based traffic accidents were recorded in Mexico City, and governmental data reports that around 3,500 people die every day on the roads. In urban and suburban areas of Mexico only in 2016 there is a record of 4,559 people dead and 97,614 people injured; for traffic accidents.

At present, systems implemented in automobiles are developed to reduce car accidents, through different types of sensors that allow quantifying variables such as the distance of objects that are in the vicinity of a car. This allows the car to make decisions to control its speed and direction among other variables. These vehicles are called autonomous and are intended to provide assistance to the driver. Due to the great concurrence in the cities, autonomous systems are currently not prepared to make a decision according to each circumstance that may arise, they also have to process a large amount of information simultaneously from different sensors. So the development of software to solve these problems is in continuous development.

An image can represent a landscape, a caricature, a page of text, the face of a person, the map of Mexico, the culture of a country, a sunset, a sunrise and endless scenarios.

The digital image is the representation of that information in a numerical matrix of intensity samples that are reflected or transmitted through the objects and that is acquired through a CCD sensor. A CCD sensor contains a series of small light-sensitive diodes that convert light into electrical charges (photons in electrons) each CCD diode captures each element that makes up the image (Pixel).

Through the images you get a lot of information about the environment as colors and shapes. Currently it is possible to estimate the depth of objects that appear in a captured scene. As the human visual system does, through stereoscopic vision (Fig. 1), we have a perception of three-dimensionality due to the difference between the images captured by the right eye and the images captured by the eye left in this way the human being is able to appreciate different distances and volumes in the environment.

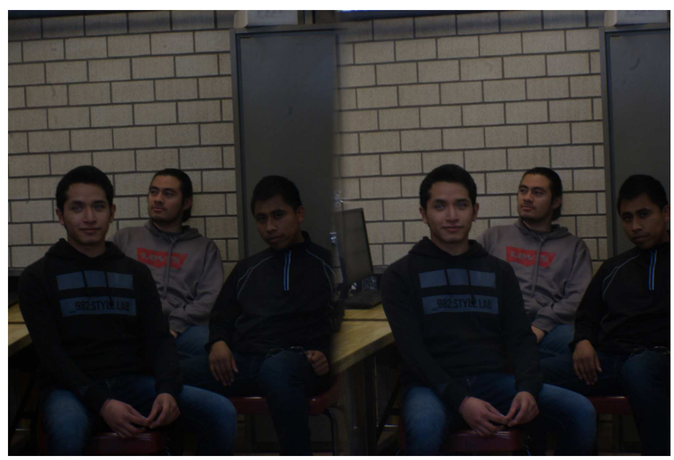

Fig. 1. Stereoscopic view.

One way to estimate the depth of each of the points in an image is by calculating the disparity, which is to measure the difference in position between two images captured by two cameras with similar characteristics, this difference or disparity is inversely proportional to the depth in the scene. To estimate the disparity, it is assumed that the scene is static, that is, that the visible objects do not change their position, nor suffer deformations. Once the disparity is estimated using one of the stereo images as a reference, an image with these values is generated, which is called the disparity map. The depth can also be quantified by means of the structured light technique that consists of projecting a burst of infrared beams of light that are captured by an infrared camera, and the depth is estimated by the distortion suffered by the projected light beams [1]

At present there are depth sensors like the one we use in this project that provides a disparity map, using the structured light technique Fig. 2.

Through image processing, algorithms implemented in autonomous navigation systems are designed, for example, object detection algorithms, pedestrian detection algorithms, traffic signal interpretation algorithms, algorithms for tracking single-lane lines, algorithms for measure the depth of the objects present in a scene, among others.

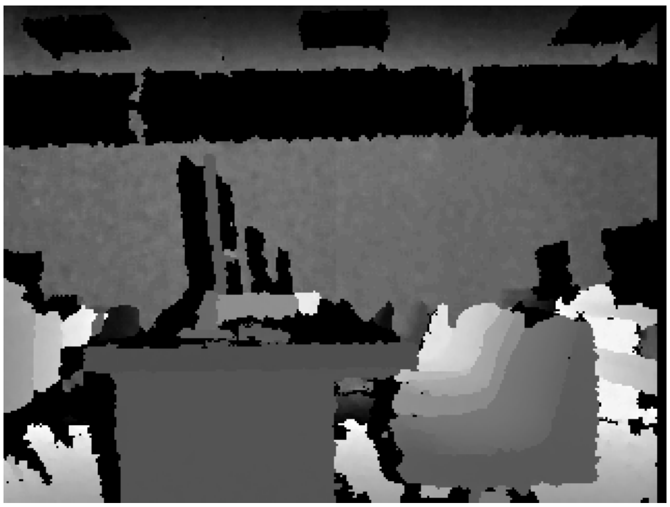

Fig. 2. Disparity map obtained using the ASUS Xtion Pro sensor.

This project presents an object detection algorithm to increase the autonomy of a mobile system, decreasing the probability of colliding with an object in the environment. The parameter to consider is the depth, which gives us the disparity map.

The depth sensor used in this project processes an image of the environment in which the autonomous system is located and as a result shows a disparity map. Specifically, the sensor captures the images at an effective distance of 0.80m to 3.50m, this implies that the system can not capture objects that are between 0 cm and 80 cm away from the sensor. As a solution to this limitation, an ultrasonic sensor is adhered to the system in the front of the vehicle, thus the system is able to detect objects that are before the field of vision provided by the sensor. To expand the effectiveness of this solution, a greater number of ultrasonic sensors is required to have a greater coverage of the environment.

Imaging is done when a sensor records the radiation that has interacted with physical objects. There are different models applied in the technology for the formation of images. One of these models is the geometric model that describes how three dimensions are projected in two. This means that it is possible to quantify the depth between the lens of a camera and the objects that appear in a captured scene.

For the development of computer vision algorithms in 3D, techniques such as monocular vision, stereoscopic vision consisting of an arrangement of two cameras with similar characteristics that take into account the epipolar geometry are used. These techniques aim to generate a map of disparity, disparity information is the process of equalization, where moves from the use of large disparities in areas of low resolution to the use of small disparities in high resolution areas.

This work is divided in another five sections: Related works, Theoretical Framework of Autonomous vehicles, Autonomous ground navigation system, Experimental Results and Conclusions.

II. Related Works

The car covers an important need in today's society, the need that people have to travel, and is that every day the individuals of a society have to move to a large number of destinations such as workplaces, schools, hospitals, even have the need to move to other cities. Because in the transfer the safety of people is at risk, automotive systems are constantly evolving each of its subsystems. And is that since its inception seeks that cars provide greater comfort and safety to the driver.

It mentions the electric ignition, ABS braking system, electronic injection, air bags, navigation system and power steering as the main improvements that have been integrated into a car. Currently, the trend is to develop smart cars with the ability to navigate, brake, avoid collisions and hazards on their own.

Artificial vision is an area of artificial intelligence that has a large part in the development of intelligent cars because it allows analyzing through images a large amount of information about the environment in which a system can be located. Through computer vision we can detect objects, colors and the distance at which an object is located with respect to a camera. Below are several advances of autonomous systems that use computer vision developed in the research and industry sectors.

Tesla is a company focused on the design of electric sports cars founded by Elon Musk in 2003. This company in the design of their electric cars has equipped them with an autonomous driving system called Autopilot [2]. As in the S model, Tesla offers an automatic pilot that drives at a speed commensurate with the traffic conditions, parks autonomously, has the ability to maintain or change lanes. Its system is composed of 8 cameras that provide a 360 degree view around the car, with a range of up to 250 meters, a front radar with improved processing provides additional data about the world with a redundant wavelength that allows you to see through of intense rain, haze, dust; twelve ultrasonic sensors that complete the vision of the cameras, allowing the detection of solid and soft objects. To understand this data, he uses an integrated computer running a neural network developed by Tesla for his visual processing software, sonar and radar.

This system provides a unique view of the world to which a driver could not have access, since it sees in all directions simultaneously and at wavelengths that go beyond the human senses, Fig. 3.

The Audi side assist is a system designed to warn drivers of dangerous situations when changing lane. Two radar sensors located on the rear bumper, which operate on the 24 GHz frequency, control the area parallel to the vehicle and also the rear area up to a distance of 50 meters. A high-speed processor analyzes the collected data. This system developed by Audi to help the driver to keep the vehicle in his lane. The signaling

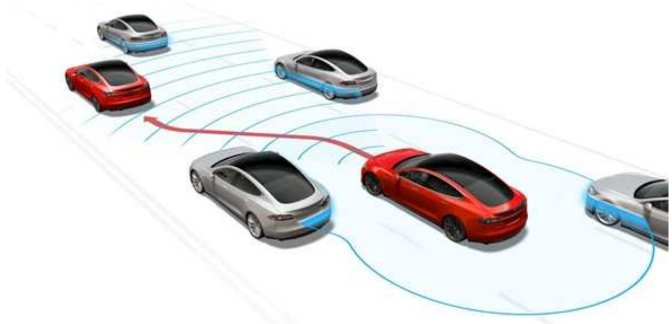

Fig. 3. Tesla Autopilot.

of the lanes is detected by a camera and defines the position of the vehicles with respect to them, Fig. 4.

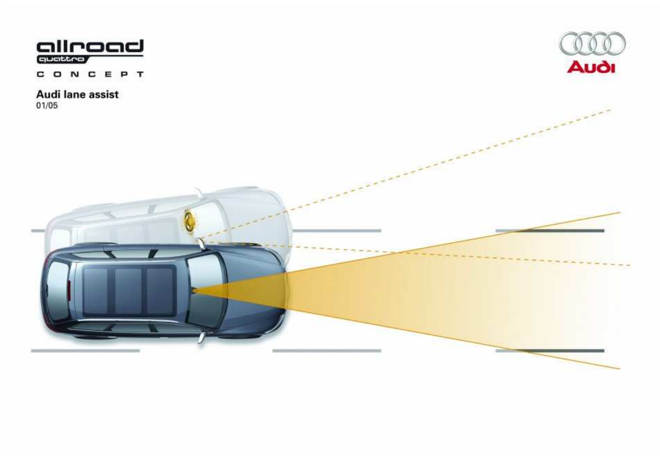

Fig. 4. Operating scheme of Audi Lane Assist.

The current interest of institutions and industry in systems capable of making driving safer has created a new line of research that seeks to address the issue of independent driving as well as computer vision, so that some of the projects developed by researchers from various organizations.

On the one hand, Combining Raspberry Pi and Arduino to form a low-cost, real-time autonomous vehicle platform proposed by Ryan Krauss in [3]. This project focuses on the creation of a low cost autonomous vehicle using Raspberry Pi, Arduino and the chassis of a juice robot. The Arduino platform handles the control laws which are executed in real time. While the raspberry pi is responsible for computer processing, a web interface and the transmission of data wirelessly. The robot has sensors to follow lines and implements a PID control to carry out its displacement, Fig. 5.

On the other hand, Real Time Obstacle Detection For Mobile Robot Navigation Using Stereo Vision is a project proposed by Sunil B. Mane and Sharan Vhanale in [4], security analysis and trends in which they perform the detection and evasion of obstacles through an algorithm. Using a passive stereoscopic Kinect camera, Fig. 6.

The fundamental idea on which they are based is to find a depth map of the image obtained by the Kinect and assign it to coordinates in the real world. The device used to achieve this is the raspberry pi 2 with which the video capture is achieved with a speed of 30 fps.

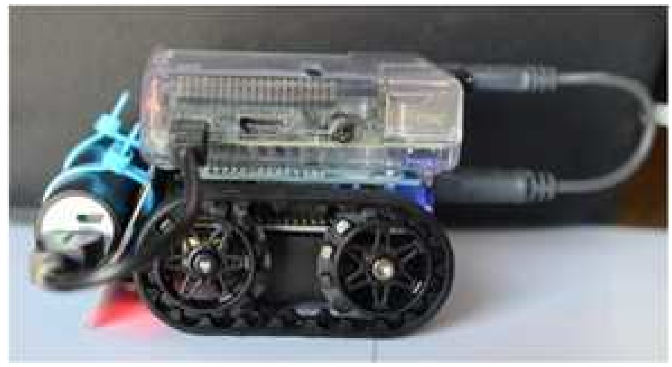

Fig. 5. Vehicle presented at the American Control Conference by Ryan Krauss in [3].

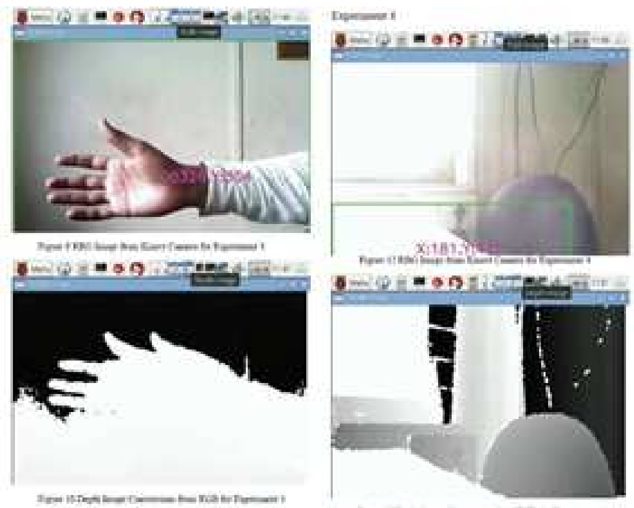

Fig. 6. Object detection using stereo vision together with Kinect [4].

III. Theoretical Framework of Autonomous Vehicles

This section will describe the hardware and software tools used for the development of the prototype as well as the general concepts necessary to facilitate its understanding.

A. ASUS Xtion Pro

Sensors such as the Xtion Pro are capable of capturing and processing gestures in real time, which makes them ideal for robots that must react to objects in motion, Xtion Pro uses PrimeSense's depth sensor technology. Table I shows the main features of this sensor. The company is known for the implementation of structured light or light coding in the infrared spectrum near light, it is an alternative to stereo camera systems that reduces costs because existing hardware can be used at low price [5].

TABLE I. ASUS Xtion Pro sensor specifications

Effective distance	0.8m to 3.5m
Field of view	58H 45V 70D
Resolution	QVGA(320x240) and VGA (640x480)

The Xtion Pro sensor only has one sensor to process the infrared light, the Xtion Pro-Live sensor includes an RGB sensor for taking photos and videos. In this way the development of solutions for implementations in computing, where a RGB camera is useful and preferable. In the development of this project, an RGB camera is not necessary because a camera gives us the requirements in the object detection solution.

Structured light is a common and economical method to obtain depth data. A light pattern is projected and recorded with a CMOS sensor. By distorting the pattern, the depth can be calculated. Most structured light sensors change the pattern several times in a capture frame to get more accurate results. Due to the use of light patterns, structured light sensors only produce adequate results in interiors and environments with controlled lighting conditions. In the case of PrimeSense sensors that produce interference patterns in the near-infrared of light can influence the performance of the sensor.

Structured light projects light patterns to generate synthetic textures, this technique makes use of a light projector and a camera, that is, a light source and a receiver of light. The laser light source is placed at a known angle with respect to the object to be illuminated and to the camera. To perform a 3D inspection of an object, a line of light is projected. The distortions in the line translate into variations in height. From here you can detach a 3D form detecting the lack or excess of material or get to do a three-dimensional reconstruction of the object. The depth information is acquired from the relative displacement of the different points of the light line by the simple line method.

B. Open NI Framework

Open NI (Open Natural Interaction) is a multilanguage, multi-platform working environment that defines the APIs for writing Natural Interaction (NI) applications. OpenNI provides a set of Apis to be implemented by the sensor devices, and a set of APIs to be implemented by the middleware components. OpenNI allows application developers to follow real (3D) scenes using data types that are calculated from the input of a sensor (for example, representation of a complete body, a matrix of the pixels in a depth map).

The OpenNI framework is an abstract layer that provides the interface for both physical devices and middleware components. The API allows several components to be enrolled in the OpenNI framework. These components are called modules, and are used to produce and process sensory data. In Fig. 7 the layers that make up OpenNI [6] are shown.

Top: represents software that implements natural interaction applications.

The middle part: Represents OpenNI, providing communication interfaces that interact with both the sensors and the middleware components, which analyze the sensor data.

The bottom part: Shows the hardware devices that capture the visual and audio elements in a scenario.

The modules supported are:

- 3D Sensor
- RGB camera
- Infrared camera
- Audio device

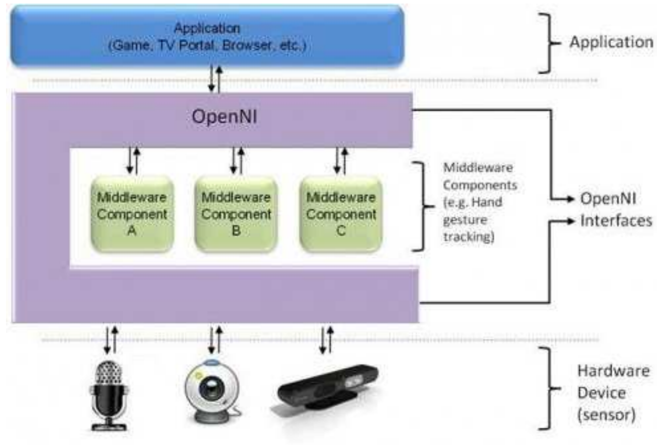

Fig. 7. Hardware devices that capture the visual and audio elements in a scenario.

C. HC-SR04 Sensor

The HC-SR04 (Fig. 8), is a sensor of distances by ultrasounds able to detect objects and calculate the distance to which it is in a range of 2 to 450 cm. The sensor works by ultrasound and contains all the electronics responsible for making the measurement. The HC-sr04 sensor consists of an emitter and an ultrasonic receiver that work at a frequency of 40KHz [7].

Fig. 8. Ultrasonic sensor HC-SR04.

The principle on which it bases its operation consists in generating a sonic wave in the emitter by means of a pulse when it hits an object, this wave is reflected towards the sensor and is recorded by the receiver. By means of the reflected wave it is possible to detect an object by measuring the time elapsed since the wave is transmitted in the air, through the *trig* pin until it is received in the echo pin. And as the speed of sound is equal to 343 m/s. With this, the HC-SR04 sensor quantifies the distance to which the sensor objects are located.

D. RASPBERRY PI3

The Raspberry Pi is a single-board computer (that is, it has all the components integrated in a single board) and a low-cost computer developed by the Raspberry Pi foundation in the United Kingdom with the aim of promoting the teaching of science in the schools Over time, the Raspberry Pi, together with the Arduino microcontroller board, have become a benchmark in the world and in the development of devices for the Internet of Things [8].

The Raspberry Pi3 has a quadcore processor much more powerful than the previous models, it also includes a Wi-Fi 802.11n and Bluetooth 4.1 which facilitates the connection to wireless networks. It includes interfaces such as an ethernet port to connect the Raspberry Pi via cable, an HDMI port that allows the connection of a monitor or TV, an audio jack output and is powered by a mini-USB connector. It includes the GPIO pins, which are used to connect the Raspberry PI to the sensors and actuators. It is also possible to connect the PiCam, the official Raspberry Pi video camera, and a monitor with DSI connection, Fig. 9.

Fig. 9. Raspberry Pi3 board.

IV. Autonomous Ground Navigation System

Explained the existing elements to develop the project and with the established work environment. In this section we will explain the operation of the system and describe in depth the methodology used for the project.

In order to implement the computer vision algorithm it was necessary to build a system whose particular characteristics in terms of hardware are the following:

- 4 moto-reducers B01 1:48.
- 4 rubber wheels.
- 2 acrylic chassis.
- 6 spacers of 3.5cm.
- Driver CON1298 (H bridge).
- Ultrasonic sensor HC-SR04.
- Asus Xtion sensor.
- Raspberry Pi3 with case.
- LiPO battery from 1500mAh to 11.1v.
- Power Bank 2200mAh with output from 5v to 1A.

Whose most relevant elements are illustrated in Fig. 10 as follows:

The functions of each element that makes up the system are described below: The structure of the prototype that includes the motors, the wheels and the acrylic chassis was chosen because its purchase was made in a kit of pieces to assemble small and according to the hardware elements that were required to be integrated into the project. This option was preferred over buying a toy with the same characteristics because in this way we would be aware of the limitations of

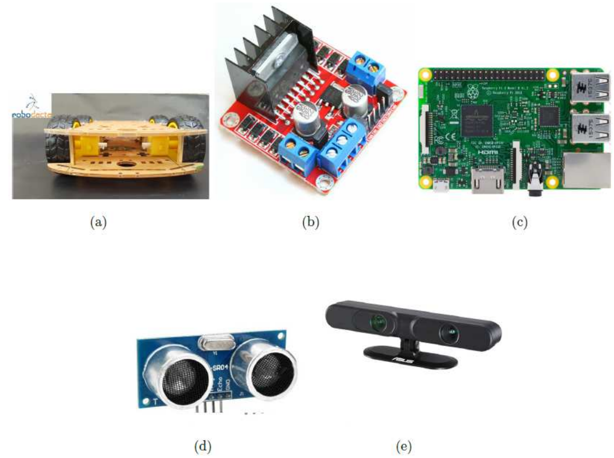

Fig. 10. Project elements, (a) Prototype chassis, (b) Driver CON1298, (c) [Raspberry Pi3, (d) Ultrasonic Sensor, and (e) Sensor Asus Xtion.

our hardware and we could integrate all the elements in a better way.

We decided to choose the CON1298 Driver for our system because it provides us with a tested and assembled H bridge, which saves us time in designing the control system for the prototype engines. The use of the driver is justified due to the need to send movement instructions from the Raspberry Pi3 to the prototype engines at different powers. This means that the instructions coming from the Raspberry Pi3 are not sent with enough power to move the motors directly, so an intermediate amplification process is needed, provided by the CON1298 Driver.

The HC-SR04 ultrasonic sensor was used in order to provide the prototype with an additional point distance measurement. This model was specifically used due to its great commercialization, its low cost and its extensive documentation. Regarding the computer vision sensor we decided on the ASUS Xtion sensor above the Kinect v1 sensor because the latter requires an additional power supply provided by an eliminator since it has additional features to those of the first mentioned sensor. The project required an infrared camera only and the Kinect sensor, in addition to providing an infrared camera, has an RGB camera, microphones arrangement and sensor adjustment motors. Therefore, its dimensions are greater than those of the ASUS Xtion sensor and its characteristics exceed those needed for the project.

For the embedded system it was decided that our best option was to use the Raspberry Pi3 development card above other options on the market such as Arduino, Beagle Bone, FreeScale, etc. Since this card, having an ARM technology processor, allows us to load a complete operating system based on a Debian distribution, as is the case with the Raspbian distribution. And with this execute more robust instructions as they are:

- Handling different USB ports.
- Higher processing power and shorter response time.
- High level programming.
- Use of concurrent programming to optimize the use of cores.

- Use of serial communication to communicate with the device.

This card when operating with an operating system eliminated the need to design the software from scratch. When using another card such as Arduino, we would have had to design all the necessary software blocks, such as reading the camera, the serial communication system to send instructions to the system, the processing of the images and the sending of control instructions to the motors. While with the Raspberry Pi3 we focus on assembling the blocks needed to perform sensor capture and serial communication to communicate with the system. Allowing us to focus on how to pose the algorithm to process the images and perform the control tasks of the prototype. top bottom The assembly of the hardware elements were arranged according to the scheme of Fig. 11.

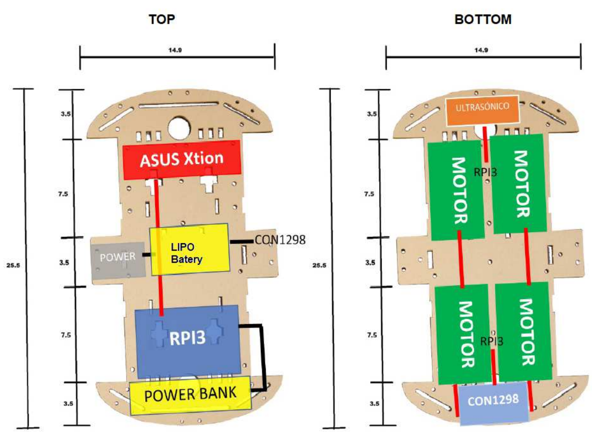

Fig. 11. Construction plan of the prototype divided into 2 parts.

Once all the hardware elements mentioned in the upper part are assembled, we obtain a system that as a whole has the following general characteristics:

- Vehicle with 4-wheel drive.
- Inclusion of a Raspberry Pi3 as an embedded control system.
- Asus Xtion as a vision system.
- Wifi connection.

The prototype of physical form is illustrated in Fig. 12 as follows:

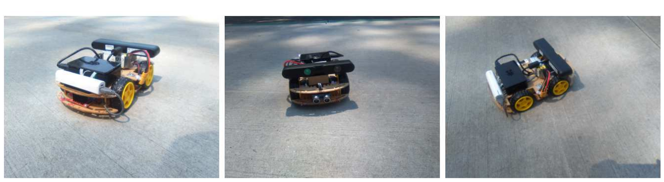

Fig. 12. Final Prototype.

The system is constituted in two fundamental parts, the one of artificial vision and the one of control of the motors. Logic and power stage respectively. Within the artificial vision a redundant ultrasonic sensor is included to have a greater security in the movements of the vehicle because the ASUS Xtion sensor works in a range of 0.8m to 3.5m so the prototype has a blind spot of 0.8m that it is compensated by the ultrasonic sensor positioned at its center, Fig. 13.

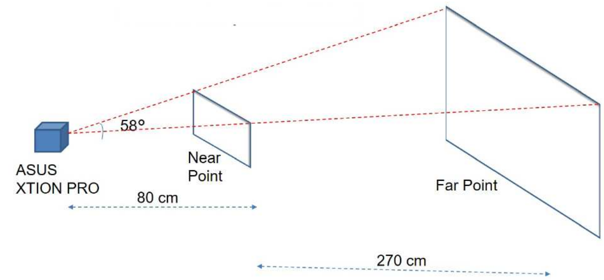

Fig. 13. Operating ranges of the ASUS Xtion sensor.

The prototype is capable of being remotely controlled from a computer thanks to the use of an access point where the embedded system connects like the computer. With which you can establish a communication between them using the SSH protocol integrated in the distributions based on Debian. This system has two independent power supplies. One feeds the control phase, while the other feeds the central computer.

The motors are connected in a series arrangement according to their position. This implies that the two motors on the left side are connected in a series, like those on the right side. So the voltage of the motors is divided into two while the current is not affected. The images captured by the ASUS Xtion sensor are processed in the following way:

1) As an initial step a capture is taken through the ASUS Xtion sensor, Which is illustrated in the image of Fig. 14(a).

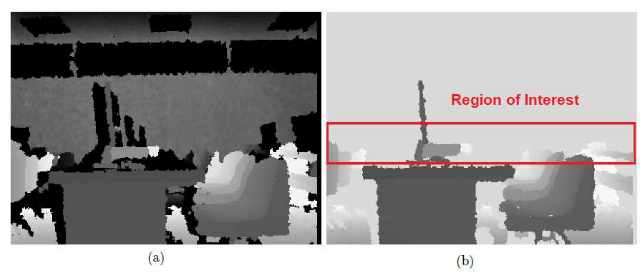

Fig. 14. Disparity map taken with (a)the ASUS Xtion and (b) thresholding.

2) The image obtained by the sensor has too many dark regions so it is necessary to use a thresholding of the image in order to eliminate the noise present in it. The result after the thresholding is shown in Fig. 14(b).
3) An image of the area of interest is obtained according to the limitations of the prototype, red box in Fig. 14(b).
4) Once the work image is obtained, it is divided into regions of interest. This is left, right and center image, Fig. 15.
5) The three images are processed by independently traversing the pixels of each one and the number of pixels less than 60 is counted. Since it was found that at 80cm from the camera the pixels are shown with

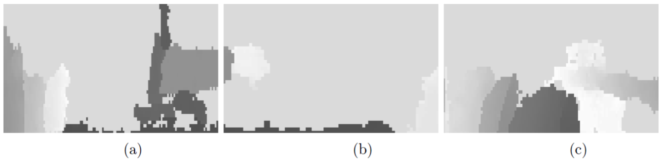

Fig. 15. Regions of interest of the image captured by the sensor, (a) Left, (b) Center, and (c) Right.

an intensity of 60. Therefore, if a pixel is registered below this value, it will indicate that the pixel is too close.

6) Once the number of pixels that are too close is quantified, validations are made to know which regions have less than 5% of the total of their pixels. Areas that have more than 5% of nearby pixels will be discarded so the prototype can not move through that area.

When comparing both images in Fig. 14. We can observe that the original capture of the sensor shows a greater depth than that obtained through the thresholding although due to the fact that the project makes the decisions with the closest images. The visualization of areas too far from the sensor becomes irrelevant since the processing of the images considers only the areas closest to the ASUS Xtion sensor.

Once the hardware elements are combined using software, the system is able to take disparity images and save them in documents with PNG extension. A methodology was designed to know the capture system and with this to know how the division should be done in areas of interest. Next, the methodology used will be explained:

1) Captures of objects at different distances were taken with the purpose of knowing at what distance the objects were visualized more clearly, Fig. 16, the objects captured at the specified distance are observed in the red rectangles. In the different captures made it was observed that at 80cm the best objects were obtained since at closer distances the objects sometimes appeared with black regions inside them, which would throw us an error in the decision of the free zones.

2) Once the distance to be used was chosen, we proceeded to delimit the dimensions of the image taken, that is to know at 80 cm of what size the frame of reference is obtained, this is illustrated in Fig. 17. These values can be obtained theoretically based on the fact that an isosceles triangle is formed between the camera and the reference frame, which can be decomposed into two right triangles, and knowing the angle of vision of the camera provided by the manufacturer, it will be possible to determine the dimensions of the reference frame theoretically.
For the assembled system it was decided to obtain the frame of reference in a practical way taking advantage of the capture of the images, This is shown in Fig. 18. In which 2 objects were placed at the ends of the capture as can be seen, then the distance between them was measured physically. This same process was carried out vertically taking as a reference the feet of a person. With which the results shown in Table II were obtained.

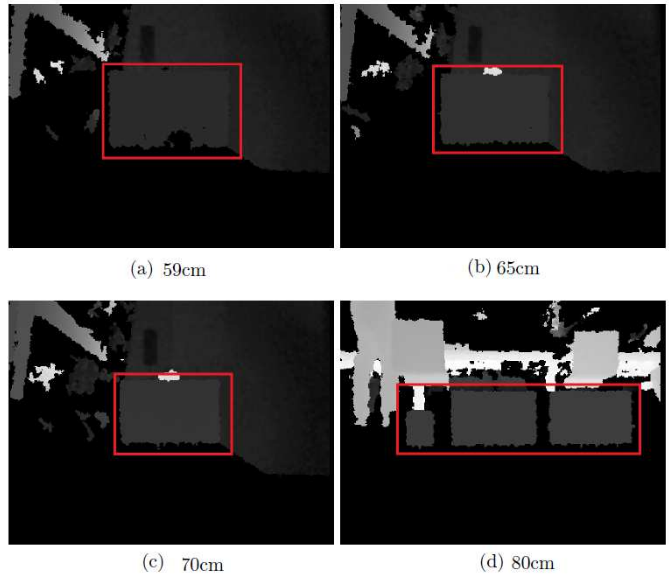

Fig. 16. Capture objects at different distances from the sensor.

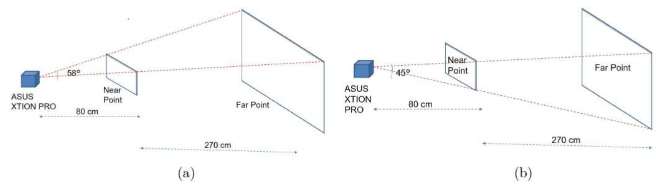

Fig. 17. Range of vision of the prototype, (a) horizontal and (b) vertical.

Fig. 18. Obtaining dimensions at 80cm.

Therefore it is concluded that at 80cm distance a pixel represents 1.17mm of the real world. To obtain the region of interest, an object of the same dimensions of the prototype is placed in front of a wall 80cm away. The image is analyzed with MatLab to determine the region that the object occupies in front of it.

TABLE II. Reference frame calculated in a practical way.

Distance(cm)	Length(cm)	Height(cm)	pixel
80	75	45cm	1.17mm
152	158		2.46mm

The object is not visually easy to identify due to its thickness but using a computer program the depth difference exists. Based on these parameters, the regions of interest described in previous sections are taken.

V. Experimental Results

The autonomous navigation system depends on the depth sensor, which integrates an infrared camera whose performance depends on the intensity of light in an environment. Below are three experiments that indicate under what conditions the functionality of the system is acceptable. Based on these experiments is that the future work of the work is considered.

A. Environment with Illumination greater than 850 lux

For this experiment the prototype was placed on a concrete slab with various obstacles with the sunlight directly incident on it and a measurement of the lux of the medium was made, which gave a measurement of 850 lux approximately at 12:00 noon. The algorithm was put into operation and it was observed that the vehicle did not detect the objects and crashed against them. In this environment the disparity map that is obtained is shown in Fig. 19.

Based on Fig. 19, it is concluded that the burst of light beams projected by the depth sensor is in the same range of frequencies that sunlight provides, due to this the light is interfered with and the image obtained is in the white spectrum. This phenomenon is similar to that experienced by a person when he tries to look directly at the sun, without any type of protective glasses.

(a) (b)

Fig. 19. Illuminated image of 850lux, (a) Real capture image, and (b) Map of disparity obtained from Fig. 19 (a).

B. Environment with Lighting less than 10 lux

In this test the prototype was placed in a roofed environment with transparent sheets which let in enough light but without the direct incidence of the sun. The algorithm loaded to the embedded system was executed and it was observed how the prototype evaded the obstacles in a precise way. In this controlled light environment, when the illumination does not exceed 10 lux, the disparity map that is obtained is shown in Fig. 20.

In the disparity map obtained, a thresholding of the objects that appear in the scene captured by the depth sensor is observed. This parameter allows the functionality of the system to be effective because it indicates that the objects that appear in a scene are identified.

(a) (b)

Fig. 20. Illuminated image of 10lux, (a) Real capture image, and (b) Map of disparity obtained from Fig. 20 (a).

C. Environment with Null Lighting

For this experiment the prototype was placed in a room with very little lighting in which the lux meter showed a measurement of 0 lux in the room. The program loaded in the embedded system was started and the disparity map of Fig. 21 was obtained.

In a nocturnal environment you get the following disparity map. As seen in the image the prototype can recognize the objects that are in the room so the system works properly in dark environments.

(a) (b)

Fig. 21. Illuminated image of 0lux, (a) Real capture image, and (b) Map of disparity obtained from Fig. 21(a).

VI. Conclusions and Future Work

An autonomous navigation device was designed making use of artificial vision which is capable of making decisions regarding its direction making use of the capture of a disparity map and processing the images. The environment of autonomous vehicles and its history was analyzed in order to know its historical context in 2019, to know what projects exist and what tools are being integrated into commercial projects such as research to identify which areas could be worked on.

The hardware and software elements that were used for the project were defined, considering their characteristics in order to understand their operation and to know how they should be assembled within the system.

The methodology used in the project was analyzed with the purpose of documenting the advances and procedures implemented within it for future reference. Within this analysis it was concluded that the prototype is not capable of functioning in sunny environments because the sun's rays cancel out the infrared lights emitted by the ASUS Xtion sensor, which gives false-positive results in the processing of the images.

Although the project is fully functional and complies with the general objective stated in the Introduction of the document, there are still several aspects that can be improved in order to obtain an improved version of it. For example the design of a communication system: a communication system based on IoT to link several prototypes and share information about their environment in order to know if there are possible collision risks between them can be designed.

Also, an analysis on the power stage: In-depth analysis of the power stage is proposed, that is to employ people specialized in control which can better model the power system and, if necessary, execute modifications in the hardware with in order to optimize the response of it. Similarly, better characterize the hardware elements to know all the limitations of this.

With the development of this project it was possible to learn about different pieces of hardware and software tools which were combined to establish a design platform on which one can work to test different computer vision algorithms, and even artificial intelligence with the purpose that students and researchers of the National Polytechnic Institute of Mexico (IPN) or other educational institutions can improve saving time in the design of the unified platform. With this tool raises a raid by Mexico and the IPN in the world of autonomous vehicles which are beginning to be the new line of research to be exploited worldwide.

Acknowledgment

This article is supported by National Polytechnic Institute (Instituto Poliécnico Nacional) of Mexico by means of Project No. 20190046 granted by Secretariat of Research and Postgraduate(Secretería de Investigación y Posgrado), National Council of Science and Technology of Mexico (CONACyT). The research described in this work was carried out at the Superior School of Mechanical and Electrical Engeniering (Escuela Superior de Ingeniería Mecánica y Eléctrica), Campus Zacatenco. It should be noted that the results of this work were carried out by Bachelor Degree students Jonathan Abraham Aparicio Osorio and César Leyva González.

References

[1] C. A. C. Solís and J. E. V. Hernández, "Algoritmo de conversión para videos teatrales," Ph.D. dissertation, Instituto Politécnico Nacional, Marzo 2015.

[2] A. M. Services, "Sistemas de asistencia audi," http://prensa.audi.es/2009/08/10/sistemas-de-asistencia-audi-eficiencia/-y-seguridad-al-servicio-del-conductor/, 2009. [Online]. Available: http://prensa.audi.es/2009/08/10/sistemas-de-asistencia-audi-eficiencia/-y-seguridad-al-servicio-del-conductor/

[3] R. Krauss, "Combining raspberry pi and arduino to form a low-cost, real-time autonomous vehicle platform," in *2016 American Control Conference (ACC)*, July 2016, pp. 6628–6633.

[4] S. B. Mane and S. Vhanale, "Real time obstacle detection for mobile robot navigation using stereo vision," in *2016 International Conference on Computing, Analytics and Security Trends (CAST)*, Dec 2016, pp. 637–642.

[5] D. M. Swoboda, "A comprehensive characterization of the asus xtion pro depth sensor," http://www.robo4you.at/wp-content/uploads/2014/11/A-Comprehensive-Characterization-of-the-Asus-Xtion-Pro-depth-sensor.pdf, 2014. [Online]. Available: http://www.robo4you.at/wp-content/uploads/2014/11/A-Comprehensive-Characterization-of-the-Asus-Xtion-Pro-depth-sensor.pdf

[6] PrimeSense, "Openni user guide."

[7] ElectroniLAB, "Sensor distancia," http://web.archive.org/web/20080207010024/http://www.808multimedia.com/winnt/kernel.htm. [Online]. Available: https://electronilab.co/tienda/sensor-de-distancia-de-ultrasonido-hc-sr04/

[8] F. Fábregas, *Aprender Raspberry Pi con 100 ejercicios prácticos*, Marcombo, Ed., 2016.