Environment Exploration Using an Active Vision Sensor

Arnaud Clérentin - Claude Pégard - Cyril Drocourt
Groupe de Recherche sur la Commande et l’Analyse des Systèmes (GRACSY)
Université de Picardie Jules-Verne - 80000 Amiens, France.

Abstract
In this paper, an omnidirectional range sensor is reported. This active vision sensor combines a CCD camera and a laser diode. We use two methods to obtain the depth of the scene: a calibration method and a least square method. We describe the prototype we made. Experimental results are presented. A comparative test shows that this sensor seems to be as accurate as a laser telemeter but less sensitive to non-alignment. Its results are better than an ultrasonic sensor. Finally, we compare three segmentation algorithms and their results on the set of points given by the sensor.

1 Introduction.

When a robot moves autonomously in a real environment, a sensor system is necessary with which the robot can recognize its own position. Estimating this position can be solved by two different methods: relative localization which uses internal sensor data and absolute localization which is based on external sensor. Internal sensors are not sufficient because errors are accumulated. In order to correct for these growing errors, it is necessary to use external sensors to update the robot position.

The solution proposed in this paper is an active external sensor which combines a laser source and a CCD camera.

Therefore, we present, in a first part, active triangulation range finders. In a second part, we explain a calibration method of this sensor. In a third part, we present a prototype of a triangulation range finder that uses a laser spot. We also describe a simple procedure to obtain depth with a least square method. Then, we compare the results of our sensor to two others sensors (a laser telemeter and an ultrasonic sensor) in order to emphasize its straight points and its applications on mobile robots. The last part is dedicated to the depth map segmentation.

2 Active triangulation range finders.

2.1 Introduction.

This method consists in projecting on the scene a visible light with known pattern geometry. A camera images the illuminated scene with a given parallax. The desired 3D-information can be deduced from the position of the imaged laser point and the lateral distance between the projector and the camera. These sensors are used to obtain depth maps, but also to digitize objects [5][9], to avoid obstacles [13][7] or to supervise industrial processes [11].

If the light source is a spot, the acquisition time essentially depends on the scanning time. This time is reduced when the light source produces one or several planes of light. In this article, we use a laser spot.

2.2 The principle of the active triangulation method.

![Figure 1. The geometric model of a triangulation system.](image)

Figure 1 shows the principle of active triangulation systems. The laser beam intersects the landmark \( M_1 \) in the point \( P_1 \). This point is projected on the retinal plane through the focal point \( F \) to a point \( u_1 \). A landmark \( M_2 \), located at an other distance, generates a point \( u_2 \). We see that the position of the points \( u_1 \) and \( u_2 \) is dependant on the distance of the landmarks \( M_1 \) and \( M_2 \). In other words, the distance of the landmark or the object \( M_i \) can be determined from the position of the point \( u_i \).

3 Calibration [4].

3.1 Introduction.

A method to obtain 3D coordinates of an environment from a triangulation sensor is to determine a geometrical model of this sensor. This stage is called calibration.
For an active triangulation sensor, the calibration can be split in three stages: camera calibration, laser beam calibration and fusion of the two calibrations to get the global model of the sensor.

3.2. The camera model.

Modeling a camera is based on the definition of a set of parameters that approximate the behavior of the physical sensor to a geometrical model. This geometrical model consists of a plane $P$ called the retinal plane in which the image is formed through a perspective projection. A point $F$ called the focal point or optical center is placed at a fixed distance from the retinal plane; this distance is called focal length. All the light beams reflected by the object pass through this point, forming a perspective image of the scene in the retinal plane (see Figure 2).

Two kinds of parameters model a camera:
- The intrinsic parameters, which model the internal geometry and optical characteristics of the image sensor.
- The extrinsic parameters, which model the position and orientation of the sensor coordinate system ($F, X_c, Y_c, Z_c$) with respect to the world coordinate system ($O_w, X_w, Y_w, Z_w$).

The intrinsic parameters are:
- The focal length $f$, which is the distance in mm from the retinal plane to the optical center $F$.
- The conversion parameters $k_h$ et $k_v$, which relate the horizontal and vertical adjustment to go from the camera coordinate system (in mm) located in the retinal plane to the image coordinate system located in the retinal plane and expressed in pixels.
- The position ($u_0, v_0$) of the image center defined by the projection of the focal point $F$ on the image plane, expressed in pixels.
- Possibly, lens distortion coefficients $k_i$.

The extrinsic parameters are the elements of a translation vector $T$ and a rotation matrix $R$.

![Figure 2. The camera model and the coordinate systems.](image)

3.3. The laser beam model.

We now want to determine the equation of the laser beam. This equation must be expressed with respect to the camera coordinate system. However, the position of this system is not exactly known. We only know the position of the world coordinate system and the extrinsic parameters. Therefore, the calibration method is:

- Determining the equation of the laser beam in the world coordinate system.
- Expressing this equation in the camera coordinate system with the help of the extrinsic parameters.

The laser beam belongs to the plane defined by the axes $Y_c$ and $Z_c$. Its equation is like:

$$Y_c = aZ_c + b$$

3.4. Last stage: sensor calibration.

Given:

- $(X_c, Y_c, Z_c)^T$ the coordinates of a point in the camera coordinate system.
- $(X_i, Y_i)$ the same point’s coordinates expressed in the image coordinate system.

The camera calibration gives the following equations:

$$X_i = \alpha_x + u_i + k_1 r^2(X_i - u_0)$$
$$Y_i = \alpha_y + v_i + k_1 r^2(Y_i - v_0)$$

where: $r^2 = (X_i - u_0)^2 + (Y_i - v_0)^2$

The laser beam calibration gives the following equation:

$$Y_c = aZ_c + b$$

From (1) and (2), we get the coordinates of the point illuminated by the laser:

$$X_c = 0$$
$$Y_c = aZ_c + b$$
$$Z_c = \frac{b\alpha_y}{Y_i - v_0 - k_1 r_2(Y_i - v_0) - \alpha_x a}$$

4. Sensor achievement.

The goal is to realize an omni-directionnal depth sensor. The effective measurable distance region is designated as 0.6m-5m: this distance is thought to be a sufficient distance for a mobile robot to detect obstacles and maneuver around them.

4.1. Remarks about the sensor design.

In order to obtain a maximum accuracy, all the height of the retinal plane must be used. So, we can make three remarks:

- The camera angle of view must be small, so the focal length must be important.
- The higher the distance between the camera and the laser is, the better the accuracy is.
- The laser beam must be inclined in comparison with the camera. Two options are possible:
  - The laser is inclined and the camera stays horizontal.
  - The laser beam stays horizontal and the camera is inclined. This configuration enables to have a depth map according to a plane parallel to the ground (we choose this configuration, see figure 1).

Since the laser beam is in the \((Y_c, Z_c)\) plane, we can use a linear camera. This kind of camera has a better resolution and a better acquisition speed than a classical CCD camera.

The sensor may be influenced by the illumination from the environment. So an infrared filter is used to extract the only light of the laser [1] [6].

4.2. Locating the laser spot image center.

We use, like [10] and [13], the center of gravity of the imaged spot. This enables us to achieve sub-pixel accuracy. Therefore, our algorithm is: finding the lightest point in the image and computing the center of gravity on a 5x5 pixels square around this point.

4.3. Obtaining the depth by a least square method.

**Principle.** An object is moved in front of the sensor. For each position of the object, we note the position in pixels of the imaged laser spot in the camera image. So we get a set of measures which associates a distance and an image position. A least square estimation of this set gives the characteristic graph of the sensor.

**Results.** We approximate the set of points we get by a six order polynomial, see Figure 3.

![Figure 3. Sensor characteristic graph.](image)

To estimate the error of this estimation, we can compare, for each calibration point, the distance computed by the least square polynomial with the real distance. We also see that the error is weak (mean 17mm, max. 40mm). That shows that the least square polynomial is a good approximation of the set of calibration points.

Another characteristic we can calculate is the sensor resolution, that is the distance represented by one pixel. For that, we compute, for a set of points, the distance \(d_1\) corresponding to a height of \(h\) pixels and the distance \(d_2\) corresponding to a height of \(h+1\) pixels. The resolution is equal to \(d_1-d_2\) for the distance \(d_1\). We remark a major default common to all triangulation systems: reduced precision with increasing range.

**Conclusion.** The first advantage of this method is its simplicity: it needs only a set of calibration points and a least square method application. Besides, this method is accurate.

4.4. Obtaining the depth by sensor calibration.

4.4.1. Classical calibration.

We here apply the calibration procedure we explain in paragraph three. We get the characteristic graph shown in Figure 4. Besides, we compute an error graph by using the calibration points of the previous paragraph (see Figure 4). We can note that the maximum error is high: 20cm at a distance of 5m.

![Figure 4. Sensor characteristic graph. (The crosses show the calibration points of the previous paragraph) and error graph](image)

4.4.2. Corrected calibration.

In order to reduce these errors, Kanade and Furhman estimate the error graph with a least square polynomial and correct these errors with this polynomial [10]. We decide to use a similar method. Our error graph is efficiently approximated by a four-order polynomial and the corrected error graph is shown in Figure 5.

![Figure 5. Corrected error graph.](image)
4.4.3. Conclusion.

Calibration is an inaccurate method [8]. The results show this inaccuracy since the sensor makes non-negligible errors (4% at five meters). A subsequent correction is needed to finally obtain a very little error (less than 1% at five meters).

4.5. Conclusion.

We built a sensor shown Figure 6. This sensor is constructed from a laser diode (wave length: 670nm, output power: 1mW) and a CCD camera (focal length of lens: 8mm, field of view: 34°). An infrared filter is used to extract only the light of the laser. The laser spot is 21cm away from the camera.

![Figure 6. The sensor we built.](image)

We do not speak about acquisition speed. Indeed, the speed essentially depends on the acquisition card.

One of the qualities of this sensor is its low cost. In comparison with the laser telemeter, it is less expensive.

Besides, this sensor is accurate. In order to judge more precisely this criterion, we will compare in the next paragraph our sensor with two other depth sensors: a laser telemeter and an ultrasonic sensor.

5. Experimental results and comparison to two others depth sensors.

5.1. The ultrasonic sensor.

The sensor consists of one rotative ultrasonic head and a step engine. The mobile part is composed of four Polaroid ultrasonic sensors with a range of ten meters and a resolution of one centimeter. The step engine has an angular precision of 200 steps per revolution. A reduction enables to position the head with a maximum precision of 0.5 degree.

5.2. The laser telemeter.

This sensor consists of a Riegl laser telemeter mounted on a rotative platform.

5.3. The experimental protocol.

We made the comparison following two points of view:

- A quantitative point of view, in order to estimate measurement accuracy and measurement repeatability. These two terms are precisely defined by an ANSI norm [2]:
  - **Accuracy** is defined as the degree of conformance between a measurement of an observable quantity and a recognized standard or specification that indicates the true value of the quantity.
  - **Repeatability** is defined as the degree to which repeated measurements of the same quantity varies about their mean.

- A qualitative point of view, in order to judge sensor abilities in a non structured environment.

5.3.1. Quantitative evaluation.

We want to evaluate the measurement error of the sensors. So, a structured environment, whose map is known with a precision of one centimeter, has been realized like the plan of the Figure 7. Each sensor made several acquisitions of this environment at the same place: at point (0,0).

![Figure 7. Environment map.](image)

**Triangulation sensor.** The measures have been realized with an angular step of 1.56 degrees. The sensor accuracy is good: a mean error of 15 mm for the best method (i.e. corrected calibration). The repeatability is remarkable: the acquisitions give similar results.

**Laser telemeter.** The measures have been realized with an angular step of 1.74 degrees. The mean error is weak (15mm) and the repeatability is excellent (standard deviation: 9mm).

**Ultrasonic sensor.** The measures have been realized with an angular step of two degrees. They are worse than the results of the other two sensors. The drawbacks of these ultrasonic sensors (multiple reflections, large emission angle, etc.) explain these poor results. The error is high (mean error: 106). The standard deviation is important (115mm), showing poor reliability.

5.3.2. Qualitative evaluation.

Let's now test the three sensors in a room of our laboratory. They are placed at point (0,0).
Trangulation sensor.

This acquisition is accurate. The different cupboards, even these grazing oriented (referenced ① on Figure 8) and the pillar (referenced ②) clearly appear. The cupboard detection shows that the sensor is not sensitive to surface orientation (i.e. a large angle of incidence).

Laser telemeter. Cupboard ① on Figure 10 has not been detected because the angle of incidence is too important: the laser telemeter sends no measure. Indeed, laser telemeters are sensitive to surface orientation: since the angle of incidence reaches 60 degrees, the measures get erroneous. Nashashibi [12] has the same conclusions, noting however that the maximum angle of incidence to obtain a valid measure depends on the surface matter. Except for this default, the acquisition is accurate.

Ultrasonic sensor. The acquisition is coarse. Only the approximate form of the room appears. In particular, for some characteristics of the room, we clearly distinguish the emission cone that prevents their detection.

5.4. Conclusion.

The laser telemeter and the triangulation sensor give excellent results. From a qualitative point of view, it would be easy to segment their depth maps. The laser telemeter is slightly more accurate than the triangulation sensor, but, on the other hand, it is sensitive to surface orientation.

The ultrasonic sensor gives coarse maps: for example, segmentation would not be easily. It seems more adapted for close perception.


Our aim is to extract line segments from the set of points given by the sensor. We have developed three segmentation algorithms and we have compared their performances.

6.1. Hough transform [3].

This method consists in constructing a repartition table. Each cell of this array represents a possible line. We get the segmentation by counting the number of points in each cell. This algorithm needs a long computation time. A same set of points can generate nearly aligned segments.

6.2. Duda-Hart recursive algorithm.

This method consists in dividing a set of point in comparison with the maximal distance of a point from the current segment. Given [Pᵢ, Pⱼ] as the current segment, we first search for point Pᵏ whose distance dstₘₐₓ from [Pᵢ, Pⱼ] is maximal. If dstₘₐₓ is higher than a given threshold, [Pᵢ, Pⱼ] is then divided in two segments (see figure 11). The same method is applied to [Pᵢ, Pᵏ] and [Pᵏ, Pⱼ]. The algorithm stops when no more maximal distances higher than the threshold can be found.

6.3. Iterative algorithm [1].

In this method from Wall and Danielsson, the points are processed sequentially. They are integrated in the current segment if they verify a criterion of cumulative error. This criterion is updated whenever a new point is added to the segment.

Let be Pᵢ the initial point of the current segment.

For a new point Pₖ, we get a cumulative error which is a rectangular surface S(Pₖ) :
- whose length is the distance d(Pᵢ,Pₖ₋₁) between Pᵢ and Pₖ₋₁
- whose height is the distance dₘₚ(Pₖ, (Pᵢ,Pₖ₋₁)) of point Pₖ from the line (Pᵢ,Pₖ₋₁), see Figure 10.

So, S(Pₖ) = (xₖ - xᵢ).((yₖ - yᵢ) - (yᵢ - yⱼ)).(xᵢ - xⱼ) -

We get Erₖ: Erₖ = Erₖ₋₁ + S(Pₖ)

We compare Erₖ to the threshold Eₛ:

Eₛ = l×\sqrt{(xₖ - xᵢ)^2 + (yₖ - yᵢ)^2} , where l is a constant threshold.

If |Erₖ| < Eₛ, Pₖ is integrated to segment [Pᵢ,Pₖ₋₁].

This algorithm has a major advantage: it does not need the complete set of points to work. However, it has two thresholds which are not easily determined and it is, like Duda-Hart, sensitive to noise.
6.4. Comparison of the segmentation methods.

Figure 11 presents the segmentation results obtained with the same acquisition (a room of our laboratory).

First, we must remark that the corners of the room in which we made our test are rounded. So the segmentation is less accurate. We can notice that the Hough transform is less efficient than the other two algorithms. Besides, it is the longest method. Therefore, we prefer either the Duda-Hart algorithm or the iterative algorithm.

In order to estimate the accuracy of these two methods, we compute, for each point, the distance between this point and its corresponding detected segment. We get table 2. Besides, we determine that the mean orientation error between the detected segments and the real segments is equal to 0.7 degrees for the two methods. These results show that our sensor gives depth maps that can be easily segmented.

<table>
<thead>
<tr>
<th>Mean distance (mm)</th>
<th>Standard deviation (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duda-Hart algorithm</td>
<td>24</td>
</tr>
<tr>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Iterative algorithm</td>
<td>29</td>
</tr>
<tr>
<td>14</td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Segmentation accuracy.

7. Conclusion.

In this paper, we have suggested and developed an active triangulation sensor with a laser beam. We have proposed and tested three calibration methods. To evaluate its performance, experiments in real environments were done. Consequently, the sensor was found to have about a mean measurement error of 15 mm which is considered an acceptable measurement error for a mobile robot. A comparative analysis with a laser telemeter and an ultrasonic sensor has been realized to conclude that our sensor is as accurate as the laser telemeter and more efficient than the ultrasonic sensor. Besides, it is non-sensitive to surface orientation. Our sensor has a low cost, is lighter and smaller than the laser telemeter. This facilitates its implementation on a mobile robot. The depth maps it provides are very accurate. Moreover, this sensor can acquire a panoramic view of the environment, in which form recognition can be done. A proximity use of the sensor in a limited angular sector is possible too.

The next work consists in matching the local map with the global map in order to extract the position of the robot.

References.