An Absolute Localization Method using a Synthetic Panoramic Image Base

Cyril Cauchois, Eric Brassart, Bruno Marhic, Cyril Drocourt

CREA (Center of Robotic, Electrotechnic and Automatic)
7, rue du moulin neuf – 80000 Amiens – FRANCE

Abstract

This paper deals with an absolute mobile robot self-localization algorithm in an artificial indoor environment. Until now, localization methods based on conical omnidirectional vision sensors uniquely used radial segments from vertical environment landmarks projection. The main motivation of this work is to show that primitives other than radial straight lines are usable with a conical vision sensor. In other words, the goal of this paper is to demonstrate that the SYCLOP sensor can be used as a vision sensor rather than a goniometric one. After a brief outline of SYCLOP geometry and the calibration process, we will show how the calibration allows us to know the omnidirectional image formation process to compute a synthetic image base. Then, we will present the localization method. Finally, some experimental results obtained with real noisy omnidirectional images acquired in an artificial indoor environment are shown.

1. Introduction

Mobile robots are increasingly more important in both industrial (navigation, exploration, surveillance, multimedia, etc.) and scientific research. Numerous research projects are aimed at increasing the autonomy of these mobile robots. To operate successfully, autonomous mobile robots must know where they are, i.e. they have to process localization. "Mobile robot self-localization" has been, and still is, a considerable subject of research. Naturally, localization techniques vary significantly according to the environment in which the robot has to navigate, prior knowledge that the robot has, the task it has to realize and the perception system of the robot. The problem is that the actual configuration (position and orientation, also called pose) always differs from the position and orientation that it is commanded to hold, because the robot is submitted to different sources of perturbations (wheel slippage, non-planarity of the ground, ...). Proprioceptive sensors (e.g. Dead-reckoning) are generally not sufficient to locate a mobile robot (location-errors accumulation), thus exteroceptive techniques have been elaborated in order to estimate accurately the robot’s configuration.

The principal difficulty of a mobile robot’s localization, from an exteroceptive sensor, is to solve the matching between the representations of the landmarks called observations provided by the robot’s sensor and the landmarks themselves. This problem could be considerably simplified if a robot’s pose (localization prediction) is given a priori. In a real situation, the absolute matching (e.g. without prediction) is quite difficult because the observations are not error-free. Moreover, the matching phase is, generally, very time consuming if one does not consider judicious strategies to reduce the combinatory aspect.

Main vision applications in mobile robotics use the classical pinhole camera model. Thus according to the lens used, the field of view is limited. Nevertheless, it is possible to enlarge the field of view by using cameras mounted in several directions, Ishiguro and al. [6]. Other applications use only one camera, with a rotation motion, in order to sweep a large space. In these two cases, algorithms of matching in successive pictures are imperative to rebuild a panoramic picture.

To get wide-angle pictures another possibility exists: omnidirectional vision sensors. Although this notion has existed for numerous years, we had to wait until the beginning of the 90’s to see their use intensified in robotic applications [13] [17] [18]. The reader will be able to get more explanations on omnidirectional vision by referring to the article of Shree K. Nayar [11]. These kinds of sensors allow us to acquire scenes with 360° field of view.

There are two major classes of omnidirectional vision systems. First of all, systems made of a mirror and a camera are called “catadioptric systems” (see applications in [1] [3] [13] [15] [17] [18]). The second one is composed of a classical camera mounted by a fish-eye lens; such mountings are called “dioptric systems” (see Cao in 1986 [4]). We focus on the first class of these sensors.
The first patent for a system using a catadioptric mechanism was registered in 1970 by D.W. Rees [15]. It was only at the beginning of the 90’s that these sensors really emerged with the utilization in robotic applications with Y. Yagi [17] [18]. In our laboratory, we have developed applications using omnidirectional sensors for robotics applications since the mid 90’s [3] [5] [13]. Our SYCLOP system, that means Conical SYstem for LOcalization and Perception, is constituted of an oriented vertical camera under a conical mirror. This kind of mirror present the advantage of providing an omnidirectional picture of an environment. In these images it is essentially the projection of the environment's vertical elements that are shown. The majority of authors using this type of mirror uniquely treat the radial features for localization, segmentation, etc. It is due to the fact that straight lines other than verticals do not project themselves following a simple mathematics model. In fact, in all these applications, the omnidirectional sensor is used as a goniometric one. Moreover, these systems present, according to Nayar, the major inconvenience of not possessing a single view point. In fact, they possess a circle of view point and therefore they generate blur pictures [1].

In this paper, we consider SYCLOP as a vision sensor. Section 2 recalls the mathematical model created for the SYCLOP sensor. Section 3 shortly recalls the protocol of calibration used to calibrate the system in its entirety. This protocol answers efficiently to constraints of markup, positioning and motives extraction of calibration points. Then we will present the SYCLOP simulator developed with the defined model. This simulator allows us to compute synthetics images closely resembling to real pictures. In section 4, we will explain the localization method based on a panoramic synthetic image base. We will show in which manner we have reduced the computational time of matching by subdividing the robot’s evolution field. Finally, experimental results with real acquisitions matched with synthetic pictures from the image base are given. We conclude our subject by a discussion on perspectives offered by this new method and how this work can be extended.

2. The SYCLOP model

2.1. Problem description

Our sensor, as the one used by Yagi in [18], is made of a conical mirror and a CCD camera with a 8.5 mm lens (see Figure 1). Nowadays, omnidirectional vision only allows us to detect all the vertical elements on a 360° domain, because they generate a set of radial straight lines converging at the center of the cone through a 2D projection.

To extend the use to the entire image, we have chosen to calibrate this sensor. First, we have to determine a mathematical model of the transformation. An object of the world will be reflected onto the conical mirror and projected onto the image plane. Figure 2a shows the sensor geometry.

The transformation contains a conical reflection that we compute with a virtual point notion. As shown in Figure 2b, the point V is the symmetrical of the point P according to the straight-line Δt. Then the point V will be projected on the CCD matrix camera. The reader will be able to get more details about the above calculations by referring to [3].

2.2. The complete model

The calculated model is based on:

- techniques of classic calibration (hard calibration) [2] [14] [16] for the intrinsic and extrinsic camera parameters determination,
- the virtual point notion for the determination of points reflected on the conical mirror.

The different stages of transformation from a real point P to its projected point onto the image plan consists of:
1. a change from the world coordinate system to the cone coordinate system,
2. a conic reflection,
3. a change from the cone coordinate system to the camera coordinate,
4. a perspective projection.

Final results of our models are given by the following equations system:

\[
\begin{align*}
\mu &= \alpha_1 r_1 V_{x1} + r_2 V_{y1} + r_3 V_{z1} + t_x + u_0 + k_x u r^2 \\
v &= \alpha_1 r_1 V_{x1} + r_2 V_{y1} + r_3 V_{z1} + t_y + v_0 + k_y v r^2 \\
V_x &= P_{x1} \left( R^2 - H^2 \left( p_{x1}^2 + p_{y1}^2 \right) + 2 R H p_{x1} p_{y1} \left( p_{x1}^2 + p_{y1}^2 \right) \right) \\
V_y &= P_{y1} \left( R^2 - H^2 \left( p_{x1}^2 + p_{y1}^2 \right) + 2 R H p_{x1} p_{y1} \left( p_{x1}^2 + p_{y1}^2 \right) \right) \\
V_z &= 2 R H \left( p_{x1}^2 + p_{y1}^2 \right) + p_{x1} p_{y1} \left( H^2 - R^2 \right) \\
\end{align*}
\]

with

\[
\begin{align*}
P_{x1} &= r_1 P_{x1} + r_2 P_{y1} + r_3 P_{z1} + i_x \\
P_{y1} &= r_1 P_{y1} + r_2 P_{x1} + r_3 P_{z1} + i_y \\
P_{z1} &= r_1 P_{z1} + r_2 P_{y1} + r_3 P_{x1} + i_z \\
\end{align*}
\]

and

\[
\begin{align*}
u' &= (u - u_0), \quad v' = (v - v_0) \quad \text{and} \quad r = \sqrt{u'^2 + v'^2}.
\end{align*}
\]

\(R\) is the radius of the conical reflector and \(H\) its height.

### 2.3. The calibration

In order to determine the set of parameters characterizing the model, we achieved an adapted calibration pattern, and a method to extract calibration points with a sub-pixel accuracy. The set of 2D and 3D calibration points gives an over determined system solved with the Levenberg-Marquardt method [12]. For a more detailed explanation, the reader is referred to [3].

Results obtained in this part are summarized in Table 1.

<table>
<thead>
<tr>
<th>(\alpha)</th>
<th>(\alpha)</th>
<th>(\bar{u})</th>
<th>(\bar{v})</th>
<th>(\bar{r})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1015.75</td>
<td>1011.16</td>
<td>384.06</td>
<td>287.79</td>
<td>-2.292 e-7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(\beta)</th>
<th>(\beta)</th>
<th>(\bar{u})</th>
<th>(\bar{v})</th>
<th>(\bar{r})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.200</td>
<td>0.503</td>
<td>0.000</td>
<td>1.814</td>
<td>-0.866</td>
</tr>
</tbody>
</table>

### 3. The simulator

With the mathematical model of our sensor SYCLOP, and with the result obtained with the calibration (Table 1), we have implemented (in C language) a SYCLOP simulator. With the help of a 3D environment map, the simulator is able to compute synthetic omnidirectional images close to real ones.

Table 2 shows some examples obtained in an indoor environment. The matching between real and synthetic images is very interesting, because even the horizontal environment’s parts are well projected onto the image.

<table>
<thead>
<tr>
<th>Example 1</th>
<th>Real image</th>
<th>Simulated image</th>
<th>Superimposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 2</td>
<td>Real image</td>
<td>Simulated image</td>
<td>Superimposition</td>
</tr>
</tbody>
</table>

### 4. The localization

In this section, we will explain the method of localization that we have implemented. Then, some experimental results with real acquisition will be shown. But first, let’s see the robot’s evolution environment. Our simulator is able to compute synthetic images in a “flat” model (no texture possibility). Furthermore, we wanted an environment without any vertical beacons in order to not be tempted to use the SYCLOP sensor like a goniometric one. Consequently, we chose to create an artificial environment comprising a lot of polygons (which act as texture) with no long vertical straight lines.
4.1. The environment

The environment of work (Figure 3) has the following measurements: 2m by 3.5m. It is composed of 5 blocks of 1.25m height and different widths. Every block is covered with a set of motives permitting a strong contour detection.

As you can note in Figure 3, we have covered each block with different motives. These motives are black on a white background in such a way that contour extraction can be easily done.

4.2. The image base

The first solution could consist in displacing the sensor everywhere in its environment in order to construct the image base. This solution has already been developed ([7] [8]) but it is not very convenient, especially if we want a thin grid. To displace the robot anywhere in the environment and to measure with accuracy the different positions is a relatively long and trying task.

The second solution will consist in using a base of synthetic pictures. With our simulator, we are able to calculate synthetic pictures close to real ones. The defect is that a sufficiently accurate 3D map of the environment is necessary, which is not the case for a real picture base. On the other hand, there is no need to displace the robot anywhere in the environment to get the image base. We only have to calculate all synthetic pictures to the different desired places.

The storage of only one picture is about 434 Ko. If we want to have a precision of half degrees, 720 pictures are necessary by position, more than 305 Mo. This solution is not foreseeable. It is preferable to only stock one picture by position (only one orientation), and compute the picture in the desired orientation. This method would certainly be very interesting, but calculations of picture rotation remain too long, even on the present computers.

All these reflections brought us to the following solution: we are going to use a synthetic picture base, but these pictures will not be omnidirectional, but panoramic. With the mathematical model of the SYCLOP sensor, and with parameters computed during the calibration, we are able to define a virtual 3D cylinder in the field of vision of the sensor. Thus, we can compute the projection of the image plan onto the virtual cylinder.

The Figure 4 shows the projection of the virtual cylinder onto the image plan. Each "square" represents an area of 10 by 10 pixels in the panoramic image.

The panoramic picture interest is double. The first advantage resides in the size of pictures. With resolutions that we fixed (1440x100), the storage of a picture requires 142 Ko. That reduced the necessary disk space by more than 3. The second advantage is the calculation of rotations. On a panoramic picture, the horizontal shift corresponds to the rotation of the picture. Calculations are simplified and accelerated.

Finally, we have subdivided the robot's evolution environment with a step of 5 centimeters. Thus, the image base is composed of 1906 synthetic panoramic images. Once we have the 3D map, the computation of the 1906 synthetics panoramic images require about 1 hour with a Pentium® III 800 MHz.

4.3. The matching

To have an image base is useful, however the major problem remains the matching. Contrary to applications that already exist [7] [8], we are going to try to match a real picture with a synthetic one. Nevertheless, our synthetic panoramic pictures are deprived of textures, and information that we possess is not on 360° but solely on five blocks distributed in the environment. Therefore, we can only match zones of pictures that correspond to projections of blocks onto the image plan.

Our first idea was to uniquely calculate 2D correlations between the real panoramic picture and synthetic ones from the base. Our panoramic pictures have a horizontal definition of 1440 pixels which permits us to calculate rotations of 0.25 degrees. So, we have 1440 2D correlations to compute for each position. This method is too gluttonous in calculation time especially as we have 1906 different positions to test.

To limit calculations, and to decrease the time of localization, we looked for a means to pass 2D correlation to a correlation of 1D signals. Elements of the scene are composed of black motives on a white background. If pictures are similar enough, the contours extraction of the real picture should be close to the contours extraction of the synthetic one. Therefore, we chose to calculate the sums in columns of binary image from contour extraction.
In this way, we will match two 1D signals while calculating their correlation. The Figure 5 shows a real picture and a synthetic picture at the same position with their segmentation and their respective sums in columns. As you can see, the two signals (e) and (f) are very similar.

4.4. The algorithm

**First stage**: By experimenting, we noted that the image base had to correspond to a grid of the environment with a step of 20 centimeters. Beyond this distance, pictures are not similar enough to ensure a correct matching. So, in our environment the initial base will be composed of 125 sums in columns. During the first stage, we select the three best correspondences according to the criteria of 1D correlation between the sum in columns of the segmentation of the real image and sums in columns of segmentation of synthetic images from the base. In this stage, solely sums in columns are used.

**Second stage**: Then, around the three retained positions, we calculate correlations with a step of 10 centimeters. It gives us 27 positions to test. For each position, we determine the best orientation with the computation of the 1D correlation with sums in columns that we weight with the 2D correlation between the real picture and synthetics ones. The same weight is given to the two correlations. The three best correspondences are selected for the continuation of the treatment. In this stage, sums in columns and images are used with the same weight.

**Third stage**: As we did previously, we are going to calculate correlations around the three retained positions, but this time with a step of 5 centimeters. Again, we have 27 positions to test. As before, we weight the correlation between sums in columns and the pictures. However, the 2D correlation between real and synthetic images has a weight of 2 against 1 for the 1D correlation between sums in columns. In this stage, correlations between pictures are major in relation to correlations between sums in columns.

The following diagram summarizes the three stages of the absolute localization algorithm:

![Diagram of the absolute localization algorithm](image)

**Figure 5**: Comparison between treatments on a real and a synthetic panoramic images.

(a) Real panoramic image  
(b) Synthetic panoramic image  
(c) Segmentation of the real panoramic image  
(d) Segmentation of the synthetic panoramic image  
(e) Sums in columns of the segmentation of real panoramic image  
(f) Sums in columns of the segmentation of synthetic panoramic image
Finally, for one absolute localization, we have a maximum of 179 tests to compute in a base that actually contains 1906 panoramic pictures. The Figure 6 shows an example of an absolute localization with all tested positions. In this example, 172 positions have been tested to estimate the robot's pose.

4.5. Experimental results

The goal of this experiment is to validate the matching method previously mentioned. Three trajectories have been implemented: two straight ones and one ovoid.

The first trajectory is straight and is 2.65 meters long. During the displacement, an image is grabbed every 5 centimeters. For this trajectory we obtain 54 pictures to treat. Figure 7 presents results obtained during the 54 absolute localizations. No bad matching is noted. The maximal gap between a supposed position, and a calculated position is 43.51 millimeters with an average gap of 22.48 millimeters. In the ideal case, all gaps would have been lower than half of the diagonal (35.35 millimeters), but the results are very satisfactory.

The second trajectory is also a straight line. It measures 2.35 meters which gives us 48 acquirements to treat. Figure 8 shows the results of matching. Once again, there were no major mistakes since the noted maximum gap is 44.96 millimeters, and the average gap is 21.66 millimeters.

Finally, the third and last trajectory is an ovoid. It measures 5.6 meters and includes 114 pictures. Figure 9 presents results of localization. This time, the maximum gap is more important since it is 56.44 millimeters and the average gap is 22.69 millimeters. But it always remained extensively satisfactory.

Figure 10 shows an example of matching between a real image (acquisition n°28 of the third trajectory) and the synthetic panoramic image base. In this example, the gap between the supposed robot's location and the position of the synthetic image associated is around 35 millimeters. On images Figure 10a and Figure 10b, we can see this difference, especially on the left of images. This is due to lighting conditions of the real environment.

As always with SYCLOP, the orientation of the robot is estimated with a high precision. During the first
trajectory, the supposed orientation of the robot was 106.54°. The average orientation calculated is 106.32°. In the second trajectory the supposed orientation was 64.86°, and the average of evaluations is 64.63°.

5. Conclusions and Future Works

The main motivation of this work was to show that primitives other than radial straight lines are usable with a conical vision sensor. In other word, the goal of this paper was to demonstrate that the SYCLOP sensor can be used as a vision sensor rather than a goniometric one.

The method of localization presented in this paper uses a synthetic picture base in order to achieve an absolute localization of a mobile robot equipped with an omnidirectional vision sensor. This method permits us to achieve an absolute localization of a mobile robot in a robust, accurate and reliable manner. Indeed, on the set of the three trajectories, that is to say 216 acquirements and so 216 absolute localizations, no matching mistake was committed. The most important gap recorded is 56.44 millimeters in an environment of 7 m². Furthermore, the use of sums in columns and the division of the process in three stage permit us to compute an absolute localization of the mobile robot in only 25 seconds on a Pentium® III 800 MHz.

We didn't treat the case of occlusion because knowledge about the environment does not cover the 360° of the field of view. In fact, in the synthetic panoramic images, a maximum of 50% of the image was known (only 5 blocks). Thus, we considered that the lack of knowledge between the blocks can be seen like occlusions in the images.

A disadvantage of this method is that it uses a base of synthetic images created using a three-dimensional map of the environment. To ensure a very good correspondence between the real images and the synthetic ones requires a sufficiently precise map and a good calibration of the vision sensor. But this defect is not inherent in this method but with all the methods which use a base of knowledge a priori (disappearance or appearance of an object). In fact, the major disadvantage of the solution presented is that the environment of evolution of the robot is completely artificial. It would have been interesting to test our method in a more "human" environment, such as offices or corridors. Thus, we are working on the improvement of the simulator in order to be able to calculate textured synthetic pictures. Another approach will consist to "untexture" real pictures in order to have an important correlation between real pictures and synthetic pictures. Finally, in this kind of environment the vertical landmarks presence would be supplementary information that we could exploit as in the previous applications of SYCLOP. This way, we will be able to compare our method to existing methods that all require an environment with vertical landmarks [9] [10].

Another future work should try to improve the high precision of localization of the robot. Indeed, the robot's pose obtained by the method presented in this paper is an approximate one. Therefore, it is possible to refine it. To achieve it, we are thinking of extracting, in an automatic manner, a set of points. It would permit us to compute a spatial localization of the robot as in [5].

Figure 10: Results for acquisition n°28 of the third trajectory.
(a) Panoramic image of the 28th acquisition – (b) Synthetic panoramic image associated with the 28th acquisition
(c) Segmentation of the panoramic image of the 28th acquisition – (d) Segmentation of the synthetic panoramic image associated with (c).
(e) Sums in columns of the segmentation of Figure 11c – (f) Sums in columns of the segmentation of Figure 11d
Acknowledgement

This work was supported in part by “Région Picardie” under the project “Pôle DIVA” (SAAC project 00-2 - “Système Actif d’Aide à la Conduite”).

References


