Vision-based System for Self-positioning and Calibration

Luis Gracia               Josep Tornero
Department of Systems and Control Engineering
Polytechnic University of Valencia
Camino de Vera s/n, Valencia
SPAIN

Abstract: - This work presents a solution for self-positioning (i.e. heading angle and offset) a vision system with respect to a line and the calibration of the single vision system. The tools Hough transform and camera model are used to obtain the position of the camera. The calibration procedure proposed calculates several parameters, which are used in the positioned process of the vision system. Experimental results, which validate the theory, are shown for two real application: industrial and agricultural environments.

Key-Words: - Vision System, Self-Positioning, Camera Calibration, Hough transform, Wheeled Mobile Robots

1 Introduction
Due to the last technological advances in vision devices, there are more and more applications using cameras as external sensors.

Many of these applications involve movements of objects or parts to be manipulated, moving equipments or even mobile robots.

As first stage the camera has to be calibrated which implies the determination of its internal parameters. Positioning implies to know the position of the vision device or camera with respect to environment, that is the elements identified in the image.

In particular, knowing the situation of the camera with respect to a line, we can obtain the wheeled mobile robot (WMR) position referred to the line. So, a WMR would be able, through its actuators, to perform a line tracking in a self-guided manner. In all these cases, it is required to identify a line like structure (e.g. the reference line of a roadway axis, a crop row, etc.).

In this sense, next section describes several ways for positioning a mobile robot with respect to a line.

Among them, in this paper a vision-based system is considered for mobile robot application.

Then, the problem is divided into two parts: on the one hand we identify the line from a preprocessed image using Hough transform, and on the other we calculate the WMR position using the camera model as described in Sections III and IV respectively.

Moreover, several parameters (focal distance, pitch angle, altitude) of the camera required for calibration will be obtained in Section V.

As a practical validation, all the stages of the process (calibration, line identification and line tracking) have been successful tested as shown in Section VI.

Finally, Section VII of conclusions remarks the most important contributions and more outstanding issues of the research.

2 Alternatives for WMR positioning

2.1 Based on distances to reference points
This alternative consists of calculating the distances between the WMR and known reference points, for a global/local positioning of the WMR.

A very used system for a global positioning of the WMRs is the GPS ([1], [2]), nevertheless many other systems may be used for a local positioning (using ultrasonic sensors, landmarks, magnetic sensors, etc.)

In this case, for a self-guided application, we have the following drawbacks: 1) the characteristics of the WMR movement space have to be previously obtained, 2) a lot of information must be saved for a good knowledge of the environment, 3) there is no robustness if some characteristic of this environment changes.

2.2 Based on image processing
This alternative consists of using a camera (as a embarked sensor fixed on the vehicle) for characterizing the image based on processing techniques. We assume that the environment is structured in lines or curves. Then, using the camera
model, the WMR is positioned with respect to this environment.

In this case, we have the following drawbacks: 1) the image processing may have a great computational cost; 2) the positioning is local. However, this alternative presents the advantage of being quite independent of the characteristics of the vehicle depending only on the reference marks. This is definitively for us, since it allows us to work in a wide variety of environments using the same positioning system.

3 Getting the reference lines from the image plane

We apply the following steps for getting the reference lines from the image plane:

1) Identify the pixels of the image plane. These pixels belong to: a) the element that represents the reference lines (band on the floor, crop row, etc.), or b) the edges of the reference line for line tracking. We apply a filter in a) and a gradient operator in b).

The second option is more convenient when the edges are computed on a straight line and it is relatively width. Therefore, we identify two straight lines, that is very faster in computation time and, then, we interpolate them.

2) We calculate the line that better adjusts the detected concern pixels. This takes the most part of the positioning computational time. So, it is important to develop a good system for line adjustment. In this sense, there exist different methods, e.g. [3] develops classifiers based on neuronal networks. However, we will use Hough transform, very common in vision applications [4].

Hough transform is used for detecting predefined geometric functions given in terms of a number n of parameters plotting a point in a space of parameters. According to that, each point in the image produces a new geometric function in this parameter space as was described [5].

In particular, we have a line as a predefined shape, characterized by two parameters. Thus, every detected concern pixel produces a curve (in general) in the parameter two-dimension space, which are all the possible couples of values for the parameters that originate lines crossing that pixel in the image plane.

In Fig.1 we show the use of the Hough transform applied to a generic image plane. We can see that the aligned points in the image plane cross the same point in the transform plane (parameter space).

We use a discrete transform space to compute the number of curves crossing each cell, and those with higher index will provide the values of the parameters of the lines identified in the image plane. In the example, the line: $y = 0.5 \cdot x + 1.5$ is identified.

![Fig.1: Example of Hough transform](image)

In this case, we have used the distance-angle representation of the line in order to prevent singularities (i.e. infinite slopes).

The preceding technique does not work well if we have a band instead of lines. In order to prevent this, it is convenient to calculate the average cell among those that, being adyacentes, are crossed by a high number of curves.

4 Positioning Calculation

First, we describe the original camera model [6] and then we particularize it for our application.

4.1 Camera model

We show in Fig.2 the basic model of image formation.

![Fig. 2: Basic model of image formation](image)

where $f$ is the focal distance of the camera; $(x_c, y_c, z_c)$ are de coordinates of a general point with respect to the coordinate system $(X_c, Y_c, Z_c)$ attached to optical center of the objective of the camera; $(x_u, y_u)$ are the projections of the point on the image plane referred to the coordinate system $(X_u, Y_u)$.

From the geometry of Fig.2, we obtain:
Expressing \((x_u, y_u)\) in terms of number of pixel dimensions \((x_p, y_p)\):

\[
x_u = \left(x_p - C_x\right) \cdot d_x \quad \quad y_u = \left(y_p - C_y\right) \cdot d_y
\]

(2)

where \((C_x, C_y)\) is the pixel position of the center point of the image and \(d_x\), \(d_y\) are the width and height of the sensing elements.

On the other hand, according to Fig.3, the relation between the world coordinate system \((X_w, Y_w, Z_w)\) and the camera coordinate system can be represented as a homogeneous transformation matrix,

\[
\begin{bmatrix}
1 & s_y & -c_y \sin \phi & c_x \sin \phi & -c_x c_y \cos \phi & s_x \cos \phi + c_x s_y \sin \phi & t_x \\
0 & c_y & s_y \sin \phi & -c_x \sin \phi & -c_y \cos \phi & s_x \cos \phi + c_y s_y \sin \phi & t_y \\
0 & -s_y & c_y \sin \phi & c_x \sin \phi & c_y \cos \phi & s_x \cos \phi + c_y s_y \sin \phi & t_z \\
0 & 0 & 0 & 1 & 0 & 0 & 0
\end{bmatrix}
\]

(3)

Note that the yaw angle between both coordinate systems is forced to be null, since it does not provide any advantage to the calculation. The pitch and roll angles are named: \(\phi, \psi\):

Obtaining \((x_c, y_c, z_c)\) from (3) and replacing the solution in (1), we get,

\[
x_u = \frac{f \cdot (\cos \psi \cdot x_u + \sin \psi \cdot y_u + t_z)}{\sin \phi \cdot (\sin \psi \cdot x_u + \cos \psi \cdot y_u + \cos \phi \cdot z_u + t_z)}
\]

\[
y_u = \frac{f \cdot (\cos \phi \cdot x_u + \sin \phi \cdot y_u - \cos \phi \cdot z_u + t_z)}{\sin \phi \cdot (\sin \psi \cdot x_u + \cos \psi \cdot y_u - \sin \phi \cdot z_u + t_z)}
\]

(4)

In general, we have in (4) five unknown variables: \{\(tx, ty, tz, \phi, \psi\}\}, meanwhile each point \((x_{ui}, y_{ui})\) from the image, for which we know its corresponding world coordinates \((x_{wi}, y_{wi}, z_{wi})\), provides two equations. Thus, we need at least three points for positioning the camera with respect to the world.

Nevertheless, if some of \{\(tx, ty, tz, \phi, \psi\)\} is certainly known, less reference points would be required.

### 4.2 Particularization of the camera model

We depict in Fig.4 and Fig.5 the top and side view of the camera, reference line and coordinate systems.

![Fig.4: Top view of the camera on the WMR and reference line.](image)

![Fig.5: Side view of the camera and reference line](image)

The new variables and parameters are: \(v\): altitude (constant) of the camera objective with respect to the reference line plane; \(P'\): point, attached to the WMR, which tracks the reference line; \(P\): midpoint between the two fixed wheels; \(h\): distance between \(P'\) and the reference line.

As there is no restriction, we situate the world coordinate system with its axe \(Y_w\) over the reference line (the selected orientation states \(Yu\) from top to down in the detected image). Therefore, the points of the reference line (and others in the same plane) have:

\[
z_w = 0
\]

(5)

Moreover, we select that origin of the world coordinate system that fulfills the condition:
This origin is placed in the cross point between the reference line and a line that, being parallel to $X_c$, starts at the cross point between $Z_c$ and the plane $zw=0$. Nevertheless, we will not need to calculate it.

As $v$ and $\phi$ are constants, and supposed known, we can obtain:

$$t_z = \frac{v}{\cos(\phi)}$$

(7)

Now, we consider (see Fig.6) points on the reference line or on parallel lines separated a distance (supposed known) $\pm d$. So, we have:

$$x_w = n \cdot d \quad \text{with} \quad n = -1,0,1$$

(8)

5 Calibration Process

The calibration process is the inverse of the positioning process, i.e. we know the world coordinates of some image points and we adjust the parameters of the vision system that make possible the correspondence with image plane.

5.1 Conditions for calibration

In order to get a simple and analytical solution, for calibration, we force:

$$x = 0 \quad t_y = 0$$

(12)

The process in a roadway tracking is shown in Fig.7.

5.2 Calculation of the pitch angle of the camera

For computing the pitch angle, we go back to (4) using (12) and (5), resulting:

$$x_u = \frac{f \cdot x_w}{\sin(\phi) \cdot y_w + t_z} \quad y_u = \frac{f \cdot (-\cos(\phi) \cdot y_w + t_y)}{\sin(\phi) \cdot y_w + t_z}$$

(13)

Operating with (13):

$$x_u = \frac{x_w}{y_u - \cos(\phi) \cdot y_w + t_y}$$

(14)

Taking two points $P_i$ and $P_j$ from the image plane, and eliminating $t$, we get:

$$\cos(\phi) = \frac{x_{ui} \cdot y_{uj} - x_{uj} \cdot y_{ui}}{y_{wj} - y_{wi}}$$

(15)
Moreover, from Fig.7:

\[
\begin{align*}
    x_{w1} &= -x_w & y_{w1} &= K \\
    x_{w2} &= -x_w & y_{w2} &= K + y_w \\
    x_{w3} &= x_w & y_{w3} &= K \\
    x_{w4} &= x_w & y_{w4} &= K + y_w
\end{align*}
\]

Equation (15) can be transformed using (16), in order to compute the pitch angle \( \phi \) from the values \( \{ A, x_w, y_w, x_1, x_2 \} \), into:

\[
\cos(\phi) = A \cdot \frac{x_w}{y_w} \cdot \frac{(x_1 + x_2)}{x_1 - x_2}
\]

Note that the couples of points \( \{ P_1, P_2 \} \) and \( \{ P_2, P_4 \} \), produce an indetermination in (15).

5.3 Calculation of the focal distance of the camera

Using (12), (10a) becomes:

\[
0 = \cos(\phi) \cdot \left( \frac{x_{ui} - x_{uj}}{y_{ui} - y_{uj}} \right) + \sin(\phi) \cdot \frac{(y_{ui} \cdot x_{ui} - y_{uj} \cdot x_{uj})}{(y_{ui} - y_{uj})} \cdot \frac{f}{(y_{ui} - y_{uj})}
\]

where \( P_i \) and \( P_j \) belong to the same reference line.

Finding \( f \) from (18), we have:

\[
f = -\tan(\phi) \cdot \left( \frac{y_{ui} \cdot x_{ui} - y_{uj} \cdot x_{uj}}{(y_{ui} - y_{uj})} \right)
\]

Then, any of both combination \( \{ P_1, P_2 \} \), \( \{ P_3, P_4 \} \) transform (19), using (16), into:

\[
f = \tan(\phi) \cdot A \cdot \frac{(x_1 + x_2)}{(x_1 - x_2)}
\]

where we calculate the parameter \( f \) (focal distance of the camera) using the known values \( \{ A, x_1, x_2 \} \) and \( \{ \phi \} \), which is obtained from (17).

5.4 Calculation of the camera altitude

Using (12) as before, (10b) becomes:

\[
0 = \frac{1}{f} \cdot \sin(\phi) \cdot \left( \frac{y_{ui} \cdot x_{ui} - y_{uj} \cdot x_{uj}}{(y_{ui} - y_{uj})} \right) \cdot n \cdot d
\]

where \( P_i \) and \( P_j \) belong to the same reference line.

Finding \( v \) from (21), we have:

\[
v = n \cdot d \cdot f \cdot \cos(\phi) \cdot \left( \frac{y_{ui} \cdot y_{uj}}{(y_{ui} - y_{uj})} \right)
\]

Then, any of both combination \( \{ P_1, P_2 \} \), \( \{ P_3, P_4 \} \) transform (19), using (16) and (20), into:

\[
v = \sin(\phi) \cdot A \cdot \frac{x_w}{(x_1 - x_2)}
\]

where we calculate the parameter \( v \) (camera altitude) using the known values \( \{ A, x_w, x_1, x_2 \} \) and \( \{ \phi \} \), which is obtained from (17).

So, using the image plane information (characterized by \( x_1, x_2 \) and \( A \)) and the world situation of the points (characterized by \( x_w, y_w \)) we get the calibrated parameters \( \{ \phi, f, v \} \), when applying (17), (20) and (23).

6 Experimental results for positioning and calibration

Fig.8 shows the application interface for calibration according to previous equations.

Fig. 8: Calibration Software Interface

The obtained values from calibration are: altitude 13.4 cm, pitch angle 45.08º, focal distance 325.7 pixels (≈325.7 x pixel real dimension). The altitude and pitch angle calibrated are very close to the values measured with traditional metrical tools (13.5 cm y 45º), what validates the calibration system proposed.

In Fig.8 we can watch an image of the reference lines (thick lines) and the straight line computed.

Note that the lines are well identified, despite that the camera lens deform the reference (slightly curved instead of a straight line).

In Fig.9, three parallel and equidistant reference lines are considered as a positioning example.

Fig. 9: Detected lines in a positioning example
Fig.10 shows the positioning of the camera ($zP=0$) with respect to each line.

Then, considering the distance between lines (3.5cm), we have a maximum positioning error (with respect to the average) of 1.5º and 2.3mm. This validates the positioning system proposed.

Other kind of reference lines can be obtained from orchard plantations as shown in Fig.11. 

This positioning system is useful in many kinds of environments: roads (using the lateral bands as references); industrial factory (using a painted line as reference in a filo-guided manner); a crop field (using the crop rows as references, etc.).

On the other hand, the calibration system developed is original and it is a simple way to obtain the vision system parameters, since it uses the same image processing as in the positioning process.

The experimental results presented in the paper validate the whole procedure encouraging us to go on doing research in this area.

Acknowledgment

We thank A. Herráez for assisting in the implementation and numerical experiments described in Section VI.

References:


