

# Derivation of Localization Algorithm for Ceiling-based Localization

March 20, 2023

## Contents

### 1 Introduction

This document formulates a solution to the following problem: A robot that can move on a 2D plane (characterized by  $x$ ,  $y$ ,  $\theta$ ) has a camera that is pointed at the ceiling above. The ceiling has a regularly spaced grid of aruco markers with different ids. We wish to estimate the robot's pose based on a single camera image. We assume that we know the world coordinates of each marker.

So we know the marker's positions in the camera images and in the world, and we wish to know the robot's pose.

### 2 The Projection Equation

Assuming a pinhole camera model, the projection equation describes where a 3D point in space maps to in the 2D camera image plane.

The equation, in its bare form is as follows:

$$\tilde{p} = K^C \tilde{P}$$

where  $\tilde{p}$  is the point's representation in the image plane (as homogeneous coordinates), and  ${}^C\tilde{P}$  is the point in 3D space, represented in the camera coordinate frame.

However, we know the marker's locations in the world frame, so we add in a transform that converts the world frame coordinates into camera frame coordinates:

$$\tilde{p} = K^{C_{TW}} {}^W \tilde{P}$$

We can write this out in matrix form:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \alpha & 0 & u_0 & 0 \\ 0 & \alpha & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta & 0 & -x \\ -\sin \theta & \cos \theta & 0 & -y \\ 0 & 0 & 1 & -0.25 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ 3.3 \\ 1 \end{pmatrix}$$

u, v, w are camera pixel coordinates (w is a scaling factor),  $\alpha$  is the focal length in pixels (we assume square pixels, and  $f_x = f_y$ ),  $u_0$  and  $v_0$  are pixel coordinate offsets. 0.25 is the height of the camera.  $x_i, y_i$  are the x, y coordinates of the center of the marker with id i. Since the ceiling is approximately 3.3 meters above, we use that as the marker z coordinate.

We can remove some of the rows and columns of the matrices:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \alpha & 0 & u_0 \\ 0 & \alpha & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta & 0 & -x \\ -\sin \theta & \cos \theta & 0 & -y \\ 0 & 0 & 1 & -0.25 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \\ 3.3 \\ 1 \end{pmatrix}$$

Multiplying  ${}^C T_W$  and  ${}^W P$  we obtain  ${}^C \tilde{P}$ :

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \alpha & 0 & u_0 \\ 0 & \alpha & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_i \cos(\theta) + y_i \sin(\theta) - x \\ -x_i \sin(\theta) + y_i \cos(\theta) - y \\ 3.3 - 0.25 \end{pmatrix}$$

If we multiply this with K, we get an expression that is hard to factorize and express in the form  $Ax=b$ . If we instead multiply each side with  $K^{-1}$  we obtain:

$$K^{-1} \tilde{p} = {}^C \tilde{P}$$

Now we can rewrite  ${}^C \tilde{P}$  and take out the unknowns. We also take out the last row and represent it as a coefficient (we convert the homogeneous coordinates into 2D, regular coordinates):

$$\frac{1}{3.3 - 0.25} \begin{pmatrix} x_i & y_i & -1 & 0 \\ y_i & -x_i & 0 & -1 \end{pmatrix} \begin{pmatrix} \cos \theta \\ \sin \theta \\ x \\ y \end{pmatrix}$$

Now we can formulate the problem in the form  $b=Ax$ :

$$3.05 \cdot K_{1:2,:}^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} x_i & y_i & -1 & 0 \\ y_i & -x_i & 0 & -1 \end{pmatrix} \begin{pmatrix} \cos \theta \\ \sin \theta \\ x \\ y \end{pmatrix}$$

Where  $K_{1:2,:}^{-1}$  corresponds to the first two rows of the inverted K matrix. Since  $w$  from  $\tilde{p}$  corresponds to a constant (it can be seen that  $w=3.05$  above), we ignore it here to be able to write the linear  $Ax=b$  relation. We only need two equations that describe  $u$  and  $v$ , so we take the first two rows of the K matrix.

This gives us our problem formulation. To apply this in localization, we take a camera image and identify all markers in the image. We take the average of each marker's corner points to get the center point, which we know the world frame coordinates to. Each marker gives us two equations relating the robot  $x, y, \theta$  to the marker coordinates. We then solve all of the equations together and find the minimum-error solution using the pseudoinverse.

### 3 Clearing up

Since we have  $\cos(\theta)$  and  $\sin(\theta)$  as two separate unknowns, we need to find a single  $\theta$  value from the two. The  $\cos$  and  $\sin$  values can be in conflict with each other: We observed a  $\sim 2$  degree difference between the two in an experiment. We normalize the  $\sin(\theta)$  and  $\cos(\theta)$  values according to  $\sin^2(\theta) + \cos^2(\theta) = 1$ :

$$\sin(\theta') = \text{sign}(\sin(\theta)) \cdot \sqrt{\frac{\sin^2(\theta)}{\sin^2(\theta) + \cos^2(\theta)}}$$

$$\cos(\theta') = \text{sign}(\cos(\theta)) \cdot \sqrt{\frac{\cos^2(\theta)}{\sin^2(\theta) + \cos^2(\theta)}}$$

Then we use `atan2` to obtain the final  $\theta$  value.

We also found out that the  $x$  and  $y$  values we obtain from our solution are not rotated according to  $\theta$  (they are in the camera frame) so we do a final rotation to obtain our robot position:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \theta' & -\sin \theta' \\ \sin \theta' & \cos \theta' \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$