Camera-LiDAR/ToF Calibration via Sphere Estimation

November 13, 2023

1 Problem Statement

The aim of this document is to show how a sphere can be estimated. Both in 3D point clouds and 2D images are considered.

2 Implicit formula for a general sphere in 3D

A sphere is exactly determined by its center point $\boldsymbol{c} = [x_0, y_0, z_0^T]^T$ and radius r. The surface of the sphere is the set of points for which the Euclidean distance is r.

In the high school, we tought the following formula for the sphere:

$$(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2 = r^2.$$

A spatial point $\boldsymbol{p} = [x, y, z]^T$ is attached to the sphere of the above equation is true.

We can reformulate the equation to obtain the implicit formula as

$$x^2 - 2x_0 x + x_0^2 + y^2 - 2y_0 y + y_0^2 + z^2 - 2z_0 z + z_0^2 - r^2 = 0 \qquad (1)$$

In the university studies, matrices and vectors are frequently used. Therefore, we reformulate the equation as

$$\begin{bmatrix} x & y & z & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -x_0 \\ 0 & 1 & 0 & -y_0 \\ 0 & 0 & 1 & -z_0 \\ -x_0 & -y_0 & -z_0 & x_0^2 + y_0^2 + z_0^2 - r^2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = 0.$$

Or, in a more compact form:

$$\begin{bmatrix} \boldsymbol{p}^{\boldsymbol{T}} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & -\boldsymbol{c} \\ -\boldsymbol{c}^{\boldsymbol{T}} & \boldsymbol{c}^T \boldsymbol{c} - r^2 \end{bmatrix} \begin{bmatrix} \boldsymbol{p} \\ 1 \end{bmatrix} = 0. \qquad (2)$$

It is a quadratic formula with respect to the homogeneous form of point p.

Remark that this is a special case of qudrics. For an ellipsoid, the implicit formula is as follows:

$$\begin{bmatrix} \boldsymbol{p}^{\boldsymbol{T}} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{M} & \boldsymbol{b} \\ \boldsymbol{b}^T & \nu \end{bmatrix} \begin{bmatrix} \boldsymbol{p} \\ 1 \end{bmatrix} = 0, \tag{3}$$

where the inner 4×4 matrix is symmetric, therefore $M^T = M$.
This quadric surface is a sphere if $\boldsymbol{M} = \boldsymbol{I}$, $\boldsymbol{b} = -\boldsymbol{c}$ and $\nu = \boldsymbol{c}^T \boldsymbol{c} - r^2$.

3 Sphere center estimation in point clouds

In Equation 1, the basic implicit formula for a sphere is given. If there is a point cloud with N spatial points, denoted by $[x_i, y_i, z_i]^T$,$i = \{1, 2, \ldots, N\}$,an equation can be written for each point as

$$x_i^2 - 2x_0 x_i + x_0^2 + y_i^2 - 2y_0 y_i + y_0^2 + z_i^2 - 2z_0 z_i + z_0^2 - r^2 = 0$$

For all points, an homogeneous linear system of equations can be written as

$$\begin{bmatrix} -2x_1 & -2y_1 & -2z_1 & 1 \\ -2x_2 & -2y_2 & -2z_2 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ -2x_N & -2y_N & -2z_N & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ x_0^2 + y_0^2 + z_0^2 - r^2 \end{bmatrix} = - \begin{bmatrix} x_1^2 + y_1^2 + z_1^2 \\ x_2^2 + y_2^2 + z_2^2 \\ \vdots \\ x_N^2 + y_N^2 + z_N^2 \end{bmatrix}$$

Remark that this problem is the same as circle fitting in the 2D plane. The solution, for the overdetermined case, can be obtained by multiplying the right vector by the pseudo-inverse of the left matrix. Them the 4D vector $\boldsymbol{v} = [v(1), v(2), v(3), v(4)]^T$ is give for which the first three coordinates are the spatial location $[x_0, y_0, z_0]^T = [v(1), v(2), v(3)]^T$ of the center. The radius can be retrieved from the last coordinate as

$$r = \sqrt{x_0^2 + y_0^2 + z_0^2 - v(4)}$$

4 Projection of a sphere into an image plane.

We are given a pin-hole (perspective) camera in which the contour points of the sphere are visible. The camera is determined by its intrinsic camera matrix

$$\boldsymbol{K} = \begin{bmatrix} f_u & 0 & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

First, we assume that the contour points can be determined in the images. Let us take a pixel in the image, its 2D coordinates are denoted by $[u, v]^T$. If the coordinate system is fixed to the camera, then the spatial ray corresponding to the pixel is written as

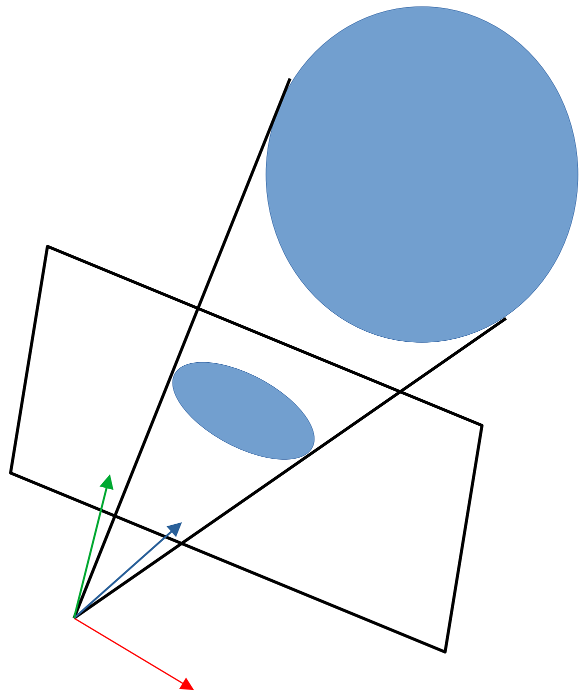

Figure 1: The controur points for a projected sphere form an ellipse in the image plane

$$\lambda \boldsymbol{K}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \lambda \boldsymbol{K}^{-1} \boldsymbol{u},$$

where parameter λ is the, usually unknown, projective depth, and $\mathbf{u}$is the homogeneous for of the pixel.

For a calibrated case, the so-called normalized pixel coordinates can be calculated. In that case, $\hat{\boldsymbol{u}} = \boldsymbol{K}^{-1} \boldsymbol{u}$.

Let us find the intersection of the ray and a quadric. For this purpose, one should substitute the ray into Equation 3:

$$\begin{bmatrix} \lambda \boldsymbol{u}^T \boldsymbol{K}^{-T} & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{M} & \boldsymbol{b} \\ \boldsymbol{b}^T & \nu \end{bmatrix} \begin{bmatrix} \lambda \boldsymbol{K}^{-1} \boldsymbol{u} \\ 1 \end{bmatrix} = 0$$

After elementary modifications:

$$\lambda^2\left(\boldsymbol{u}^T\boldsymbol{K}^{-T}\boldsymbol{M}\boldsymbol{K}^{-1}\boldsymbol{u}\right)+\lambda\left(2\boldsymbol{b}^T\boldsymbol{K}^{-1}\boldsymbol{u}\right)+\nu=0$$

This is a quaratic equation w.r.t. λ. Such kind of equations has two, one or zero real roots.

Let us examine the geometric meaningif the number if roots. There can be two, one or zero intersection point of a ray and a sphere. if the pixel corresponds to the contour, then the ray touches the sphere, in that case, only one intestecting point exists.

Algebrically, there are only one root for the quadratic equation. It is possible if and only if its determinant is zero. Therefore,

$$\left(2\boldsymbol{b}^T\boldsymbol{K}^{-1}\boldsymbol{u}\right)^2-4\left(\boldsymbol{u}^T\boldsymbol{K}^{-T}\boldsymbol{M}\boldsymbol{K}^{-1}\boldsymbol{u}\right)\nu=0.$$

For the sphere, $\boldsymbol{M}=\boldsymbol{I}$, $\boldsymbol{b}=-\boldsymbol{c}$ and $\nu=\boldsymbol{c}^T\boldsymbol{c}-r^2$. For normalized coordinates, $\hat{\boldsymbol{u}}=\boldsymbol{K}^{-1}\boldsymbol{u}$. Thus,

$$\left(-2\boldsymbol{c}^T\hat{\boldsymbol{u}}\right)^2-4\left(\hat{\boldsymbol{u}}^T\hat{\boldsymbol{u}}\right)\left(\boldsymbol{c}^T\boldsymbol{c}-r^2\right)=0.$$

After elementary modification:

$$\left(\boldsymbol{c}^T\hat{\boldsymbol{u}}\right)^2+\left(\hat{\boldsymbol{u}}^T\hat{\boldsymbol{u}}\right)\left(r^2-\boldsymbol{c}^T\boldsymbol{c}\right)=0.$$

If we introduce the homogeneous coordinates:

$$\hat{\boldsymbol{u}}=\begin{bmatrix}\hat{u}\\ \hat{v}\\ 1\end{bmatrix},$$

then

$$\left(\hat{u}x_0+\hat{v}y_0+z_0\right)^2+\left(\hat{u}^2+\hat{v}^2+1\right)\left(r^2-x_0^2-y_0^2-z_0^2\right)=0.$$

After elementary operations, it becomes

$$\hat{u}^2x_0^2+\hat{v}^2y_0^2+z_0^2+2\hat{u}\hat{v}x_0y_0+2\hat{u}x_0z_0+2\hat{v}y_0z_0+$$

$$\left(\hat{u}^2+\hat{v}^2+1\right)\left(r^2-x_0^2-y_0^2-z_0^2\right)=0.$$

This equation can be written as a conic:

$$A\hat{u}^2+B\hat{u}\hat{v}+C\hat{v}^2+D\hat{u}+E\hat{v}+F=0,$$

if

$$A=r^2-y_0^2-z_0^2, \tag{4}$$

$$B = 2x_0y_0,$$

$$C = r^2 - x_0^2 - z_0^2,$$

$$D = 2x_0z_0,$$

$$E = 2y_0z_0,$$

$$F = r^2 - x_0^2 - y_0^2.$$

Remark that this conic is an ellipse if the sphere is in front of the image.

5 Estimation of sphere centers in camera images

If at least five contour points in the images are known, them we can estimate its parameters by e.g. the Fitzgibbon method. Then A, B, C, D, E, F are known.

It is known that

$$\frac{B}{D} = \frac{y_0}{z_0} \rightarrow Bz_0 - y_0D = 0,$$

$$\frac{B}{E} = \frac{x_0}{z_0} \rightarrow Bz_0 - x_0E = 0,$$

$$\frac{D}{E} = \frac{x_0}{y_0} \rightarrow Dy_0 - Ex_0 = 0.$$

In a more compact form, it is modified as follows:

$$\begin{bmatrix} 0 & -D & B \\ -E & 0 & B \\ -E & D & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} = 0.$$

The matrix in the left is linearly dependent as its determinant is zero.

The solution is given as the null vector of the left matrix. Unfortunately, the solution is detemined only up to an unknow scale as $\alpha\left[x_0, y_0, z_0\right]^T$ also satisfied the previous equation. This in not a real surprise as the scale cannot be reconstructed from camera images.

However, if somebody knows the radius of the sphere, then the scale can be extracted. For example, from Equation 4, it can be written that

$$A = r^2 - \alpha^2 y_0^2 - \alpha^2 z_0^2.$$

Then

$$\alpha = \pm\sqrt{\frac{r^2 - A}{y_0^2 + z_0^2}}$$

The sign ambiguity can be removed as only one of the signs yield a sphere behind the camera. The other solution is behind the camera, thus that can be discarded.

6 LiDAR-camera calibration

If we have the complex task to register a camera and a Lidar, or a camera and a Time of Flight (ToF) sensor, the task is as follows:

1. Put a sphere with know radius into the scene, take a picture by the camera, record the point cloud by the LiDAR/ToF sensor.
2. Estimate the sphere center from the point cloud by the method in Section 3 .
3. Estimate the sphere center from the contour points in the image by the method in Section 5.
4. Repeat steps 1-3 at least four times.
5. Take the sphere centers, the compute the rotation and translation between camera and point cloud centers by the optimal point registration algorithm.