An Extrinsic Calibration for Radar, Camera and Lidar.

Why Calibration?

→ In order to represent sensor observations in a common reference frame, the rigid transformation (3D rotational and translation) between all sensor co-ordinates frames must be known.

Extrinsic Calibration can provide more precise estimates, as it alligns corresponding sensor measurements of the real targets directly.

proposed approach.

Steps:

a) Calibration board design is described
b) Detectors that extracts features from sane sensor data are described
c) A mathematical description of the Calibration of 2 sensors.
d) Joint Calibration of more sensors.

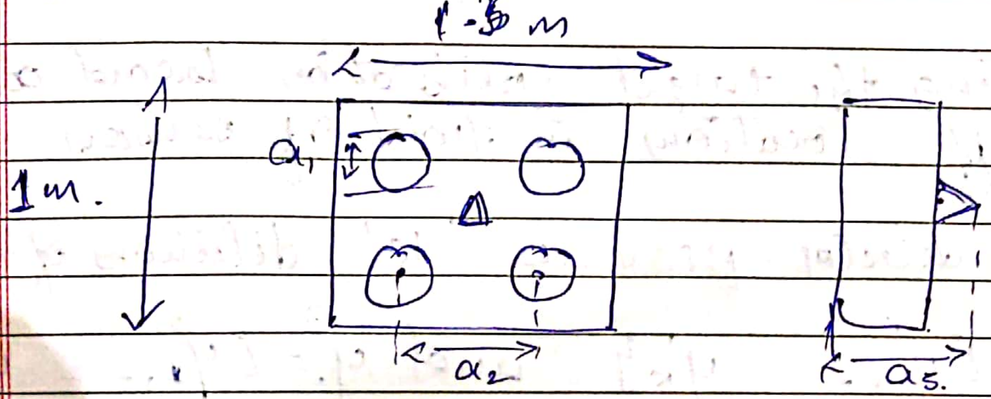
a) Calibration target design.

Calibration target should be detectable in all relevant modalities for multi-sensors calibration.

For lidar and Camera, edges and corners are features which can be detected accurately

However rectangular shape objects are difficult to localize in lidar, as a horizontal ~~planes~~ edge might not intersect with any of the lidar scan planes.

Hence used circular shapes, which can be accurately detected when intersecting with lidar beams.



Target design has 4 circular holes and a single metal tetrahedral corner reflector in the center at the back of the board to provide strong radar reflections.

a_1 = circle diameter = 15 cm
a_2 = distance between centers = 24 cm
a_3 = reflector is at 10.5 cm from front.

b) Detection of calibration target (detector)

Both the camera & lidar detectors returns the 3D locations of the 4 circles centers.
The radar returns detections in a 2-D plane and generates for each reflection a measurement in polar coordinates & Radar Cross Section (RCS) values. Of all the detections that are within the expected RCS range, the closest radar detection to the car is taken.

c) Calibration procedures:

Extrinsic calibration of 2 sensors.

Placing the target calibration board at 'K' diff locations in front of sensors.

Each detector provides 'K' detections of target,

$y^1 = \{y^1_1, \ldots y^1_K\}$ and $y^2 = \{y^2_1, \ldots y^2_K\}$.

for sensor 1 for sensor 2

These detections are relative to co-ordinate frames of sensor 1 and 2 respectively.

For camera & lidar, each detection consists of '4' 3D co-ord's of the circle centers.

i.e $y_K = (y_{K(1)}, \ldots y_{K(4)})$

A Radar has only a single location $y_{K(1)}$ for the detected tetrahedral corner reflector, which is expressed in 2D Euclidian co-ord.

i.e $y_K = (y_{K(1)})$.

Since each sensor has different field of view (FOV), the calibration target may not always be detected by all.

M_k^i → Indicates if calibration board k was detected for sensor i.

if $M_k^i = 1$ → target was detected.
if 0 " " not ".

Extrinsic calibration of 2 sensors aims to estimate the relative rigid transformation ($T^{1,2}$), which projects a point from sensor 1 onto the coordinate frame of sensor 2.

Rigid transformation ($T^{1,2}$) consists of:
1. 3×3 rotation matrix 'R'
2. 4×4 3D translation matrix 't' for homogeneous co-ordinates.

$$T^{1,2} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$$

we assume homogeneous representation in y^1 and y^2, hence each 3D points (x, y, z) is represented as 4D vector $(x, y, z, 1)$.

Parameter vector, $\theta^{1,2} = (t_x, t_y, t_z, V_x \cdot \alpha, V_y \cdot \alpha, V_z \cdot \alpha)$.
↓
To parameterize 6 DOF transformation

The rotation is expressed by axis-angle representation, as a unit vector (V_x, V_y, V_z)

For the axis of rotation, & an angle α

$$T^{1,2} = \begin{bmatrix} r_{xx} & r_{xy} & r_{xz} & t_x \\ r_{yx} & r_{yy} & r_{yz} & t_y \\ r_{zx} & r_{zy} & r_{zz} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}_{4\times4}$$

For K^{th} target location, the transformation error between camera and/or lidar detection is the total squared Euclidean distance of the 4 detected circle centers.

$$\epsilon_K(\theta^{1,2}) = \sum_{p=1}^{4} \left\| y^2_{K(p)} - T^{1,2} \cdot y^1_{K(p)} \right\|^2$$

If one of the sensor is Radar, then,
Let, $y^R_K \rightarrow$ radar measurements of target K.

$$\epsilon_K(\theta^{1,R}) = \left\| y^R_{K(1)} - p\left(T^{1,R} \cdot g(y^1_K)\right) \right\|^2$$

Here,
$g(y_K) \rightarrow$ computes the expected 3D position of the tetrahedral corner reflector given the 4 3D circle positions in detection y_K

$p(q_K) \rightarrow$ first converts 3D Euclidean point q_K to spherical co-ord (r_K, ϕ_K, φ_K), then discards the elevation angle φ_K, then converts remaining polar co-ord (r_K, ϕ_K) to their 2D euclidian co-ord equivalent.

Additionally, we enforce the constraint that the projected 3D points lie within Radar FOV.

Hence add constraint that elevation angle ψ_K for all calibration board location K are within max^m view angle $\psi_{K, max}$ of radar.

i.e $|\psi_K| - \psi_{max} \leq 0$.

$f(\theta^{1,2}) \rightarrow$ total error between all K calibration targets.

i.e $f(\theta^{1,2}) = \sum_{K=1}^{K} \mu_K^2 \cdot \mu_K^1 \cdot \epsilon_K(\theta^{1,2})$.

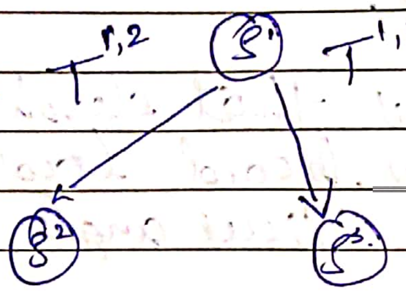
$\mu_K^2, \mu_K^1 \rightarrow$ ensures that only terms are included where the target was detected in both sensors.

The optimal calibration parameters are found by minimizing $f(\theta)$, which could be subject to θ.

d) Joint calibration with more than 2 sensors.

We Consider 3 Configurations to adapt the extrinsic calibration procedure to optimize θ.

1) Minimally Connected Pose estimation (CMCPE).



$S^i \rightarrow$ Sensors.
$T^{i,j} \rightarrow$ Co-ordinate frame transformation from Sensor i to j.

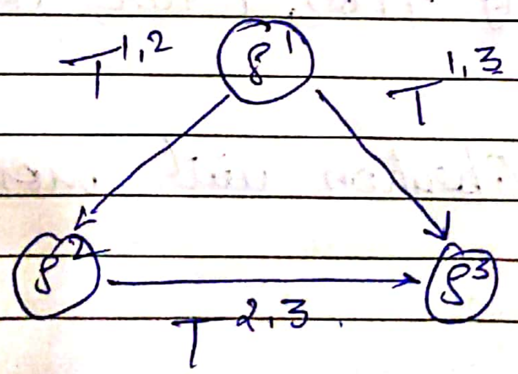
Sensors can be calibrated pairwise with respect to a selected 'reference' Sensor (S^1).

Optimization function :

$$f(\theta) = \sum_{i=2}^{N} \left[\sum_{k=1}^{K} M_k^i \cdot M_k^1 \cdot \epsilon_k(\theta^{1,i}) \right]$$

i = no. of sensors
K = no. of detections.

2) Fully Connected Pose estimation (FCPE)



we Consider optimizing transformations between all sensors at once, without a special reference sensor.

Instead of estimating N−1 transformation matrices with respect to a reference sensor, all transformation matrices between all $\binom{n}{2}$ combination of 2 sensors are estimated.

$$\therefore f(\theta) = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left[\sum_{K=1}^{K} M_K^{i} \cdot M_K^{j} \cdot \epsilon_K(\theta^{i,j}) \right]$$

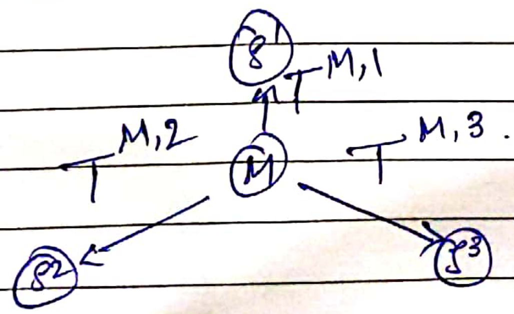
In addition, the closed loop constraint is introduced to ensure that all loops 'l' equal the Identity matrix.

i.e. $\left(T^{S_l,1} \cdot T^{S_l-1,S_l} \cdot, \ldots, T^{1,2} \right) - I = 0$

$S_l \Rightarrow$ the no. of sensor in the loop 'l'.

By adding more error terms, the optimization is potentially more robust against noisy observations from one reference sensor.

3) Pose and structure estimation



This configuration explicitly estimates the calibration board poses and ~~structure estimation~~ the observation noise of each sensor.

Objective is to estimate :

(1) unknown structure of the free poses in a fixed co-ord. frame. i.e $M = (m_1, \ldots, m_K)$

(2) Transformation from the fixed frame to each sensor i.

i.e $T^{M,i}$

Observations are considered samples from a probabilistic measurement model.

$$y^{i}_{k(p)} = T^{M,i} \cdot \hat{y}^{M}_{k} + \eta^{i}$$

zero mean gaussian noise

Mahalanobis dist

$$D^2_{\Sigma}(a,b) = [a-b]^T \cdot (\Sigma)^{-1} [a-b].$$

a & b → Vectors

Σ → Covariance.

For pose estimation, we first initialize all Σ^{i} as identity, and jointly optimize the transformations and structure.

$$E_K(\Theta^{M,i}, M) = \sum_{p=1}^{4} D_{\Sigma^i}^2 \left(y^i_{K(p)}, T^{M,i} \cdot \hat{y}^M_{K(p)} \right).$$

$$f(\Theta, M) = \sum_{i=1}^{N} \left[\sum_{K=1}^{K} \mu^i_K \cdot E_K(\Theta^{M,i}, M) \right]$$