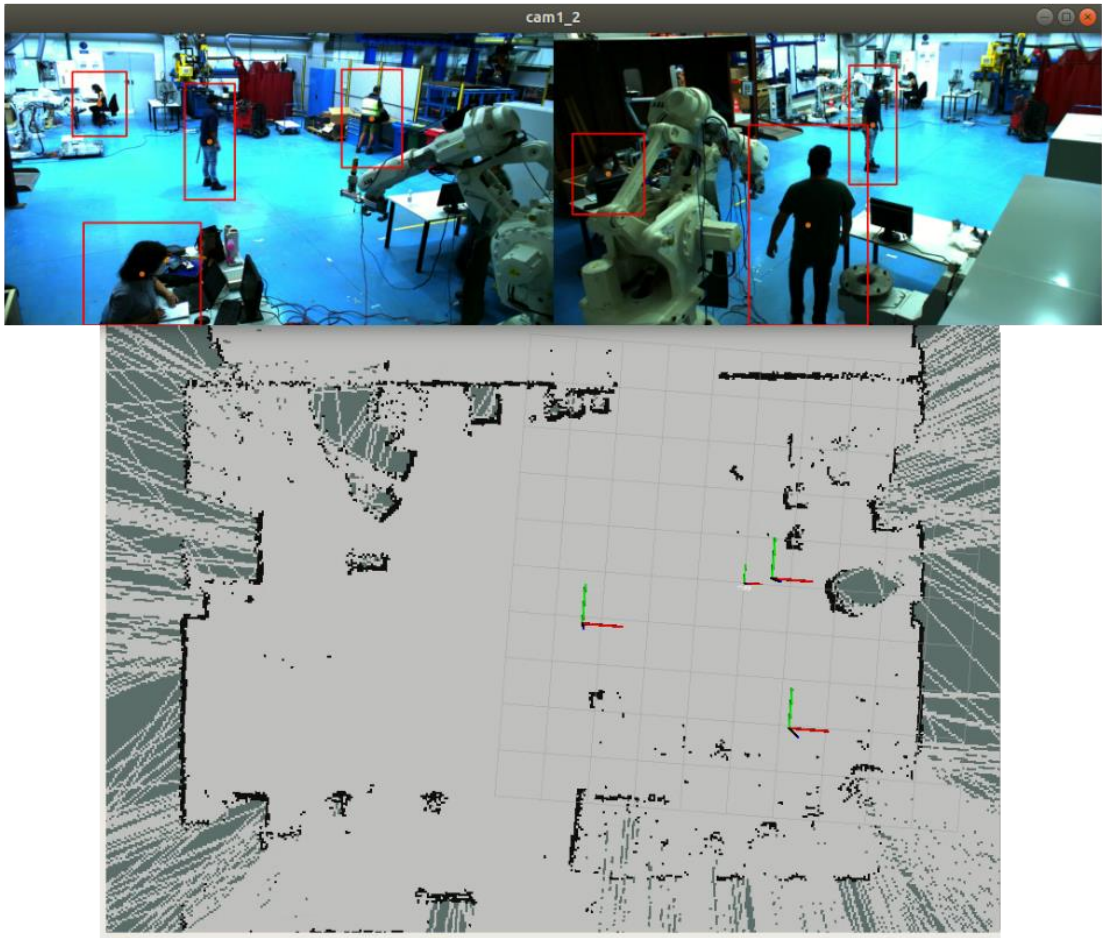
Person detection & Tracking

Overview

The Person Detection and Tracking systems is based on 2D color images from cameras installed in environment and it is intended to monitor the shared spaces between humans and robots. It allows to detect people and track them using stereo pairs, obtaining the pose of each person detected in the navigation map. It requires a metric map to update dynamically the occupancy grid, information that could be used to implement local path planning to avoid dynamically occupied places, or it could be used to know the busy areas in the shopfloor for a better deployment of mobile platforms and optimization of available resources.

Image processing is based on YOLO, used to identify persons in every image acquired from different sources. The Regions of Interest obtained with YOLO from every image are paired using a combination on geometric computation and tracking using Minimum Output Sum of Squared Error (MOSSE) for disambiguation when required. The pose obtained is translated to the occupancy grid.



Objectives

- Detect people in 2D images from environment cameras.
- Extract the pose of every person using stereo pairs.
- Track persons and translate it to occupancy grid.

Hardware requirements

At least two RGB cameras with their corresponding optics to stablish one stereo pair and a computer with average performance (GPU desirable but not mandatory). The module has been tested with IDS cameras, but it could be generalized to adapt it to any hardware. It is important to note that specific acquisition procedure is necessary to ensure the images synchronization.

Deployment

One of the essential steps in order to be able to compute 3D positions from stereo vision systems, is to previously calibrate the stereo pair. This calibration involves the intrinsic calibration of each camera and the stereo calibration of both in order to know the transformation between the two cameras (relative position).

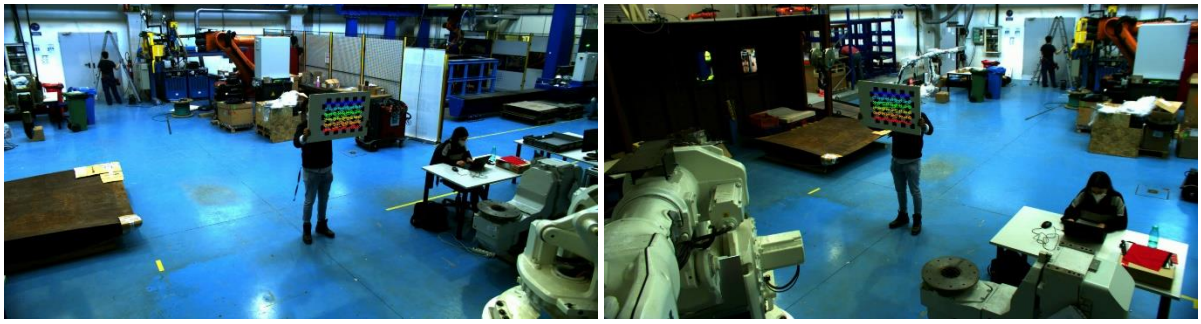


Figura 1. Images of the stereo calibration process.

Both intrinsic and extrinsic calibration have to be made with a chessboard pattern and the OpenCV library tools. In the following images several images used for calibration are shown.

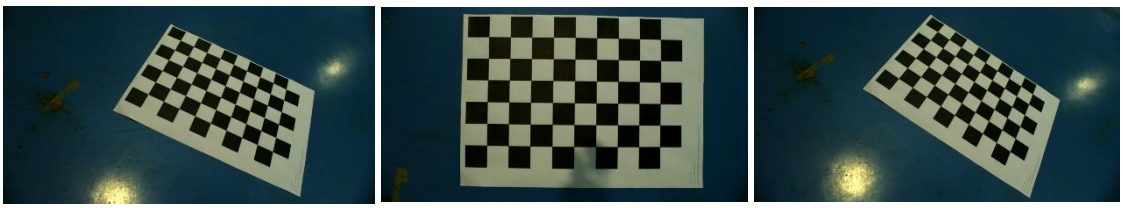


Figure 1. Example of images captured for individual calibration.

Figure 2. Example of images captured for stereo calibration.

Integration with OPIL

The module LAO_Human_Tracker is integrated in OPIL, so the persons detected will be added to the virtual representation of the environment through the FIROS bridge. To send the information about the position of the operators we are using the Local Sensing & Perception module. We have done this because the type of message we need in our application has already been defined and implemented in the Central Sensing & Perception module and in the Local SP.

We are using the same type of messages that are used to indicate the position of the robots within the map, and it is the following:

Message defined in: Link.

Topic	Message Type	Format
/worker/newObstacles	Mapupdates/NewObstacles	Header header Float64 [] Float64 []

We need to make a few minor changes within the Central SP to subscribe to the new topic and have the position of the workers considered in the map. For both the Local SP and the Central SP, the FIROS module must be parametrised to subscribe and advertise this topic respectively. Once the changes have been applied, the exchange of messages within the Central SP looks like in the picture.