Human-Aware Navigation Using Kinect-based Active Perception

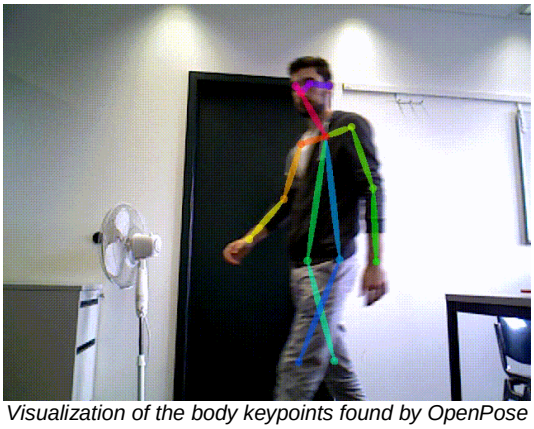
Nicolas Talabot

Professor : Alcherio Martinoli
Assistant(s) : Zeynab Talebpour

Social mobile robots are systems evolving in an environment shared with people, and therefore require human-aware navigation. A crucial aspect for this behavior is the ability to detect and track people accurately. This means estimating their positions and orientations, and inferring their speeds, and possibly goals, thanks to the tracking. It can be solved by using external perceptions systems, such as fixed cameras on walls, or on-board sensors. The latter allows for much more flexibility as no control or prior modification of the environment is needed. In this project, we were interested in using the on-board Kinect sensor of the MOnarCH robot for robust person detection and tracking, to estimate the position and orientation of people. It was done while aiming towards controlling the head motion of the robot, namely, active perception.

Initially, we were to use an already existing code, and implement it on our robot. We however realized that it was not practical as the code could not be easily adapted. Therefore, we spent time searching for other software during the semester, and the advancement of the project has then been quite limited by this.

We ended up with two detectors/trackers, namely *cob_openni2_tracker* and OpenPose. The first one is a ROS package directly made to use the Kinect to detect and track users. It was not found really robust, as it needs a calibration phase per user that can require up to a few seconds, which is not efficient in a real time scenario. Nevertheless, it is contained in a meta-package that could provide useful modules, and was thus kept. About the second, OpenPose is an independent software that can detect human keypoints on single color images. Hence, it does not directly use the depth information of the Kinect, and requires a wrapper to work with ROS. Nonetheless, it seemed more robust and promising, and as a result we decided to focus on it for the remainder of the project, despite its high computational cost.



Visualization of the body keypoints found by OpenPose

We found a ROS wrapper for OpenPose that additionally projects the keypoints onto the 3D point cloud provided by the Kinect. Using these information, we wrote a node estimating the position and orientation of a detected person given its keypoints. Moreover, as OpenPose does output a confidence score for each keypoint, we were able to provide one for the pose too, that can be taken as an uncertainty measure.

Following this, we used the little time left to assess the performance of this node in a qualitative manner, as we had not access to an exact ground truth. It was found to work quite well, even though the accuracy was not perfect. The detections appeared to work with walking people and camera motions, but further testing could be effectuated with the right equipment.

In addition, some improvements are possible, and the frame rate is still quite low because of the computational cost of the detection. The current detections are frame per frame, and a tracker still needs to be implemented, and could be used to filter the results of the detections. Finally, the active perception could be written, using the provided confidence.