Autonomous Warehouse Navigation using 3D Time-of-Flight Cameras

Marian Himstedt and Ulrich Behrje and Erik Maehle

Abstract— **This paper presents an overview of the navigation system of an autonomous forklift being developed within the publicly funded project FTF-out-of-the-box. The control of the vehicle is accomplished by means of storage requests initiated by a human operator using speech control. Instead of the commonly used 2D laser range finders we make use of Time-of-Flight (ToF) cameras for the entire vehicle navigation. We explain what parts of typical navigation stacks have to be adjusted for this type of sensor. Experimental results obtained in a warehouse environment are presented.** [1]

I. Introduction

Today automated guided vehicles are commonly used in warehouse environments. The majority of available systems are centrally planned by warehouse management systems. Within an initial setup (teach-in), technicians generate layouts of road plans which are subsequently followed by AGVs. The position estimation of the vehicle is carried out by the use of 2D laser range finders and artificial landmarks or by magnetic track guidance. In recent years there has established a notable amount of commercial systems enabling contour-based localization without the use of artificial landmarks. However, all relevant points in the environment, which are aimed to be approached by AGVs, are taught in the reference frame of the map. This imposes notable requirements on the system as tasks such as picking up pallets demand a high position accuracy. Our work differs from the common navigation systems as follows:

1) All drive requests are provided by a human operator using a speech-based HMI (decentralized system) [1].
2) The position estimation copes without setting up artificial landmarks.
3) Storage locations in high-level racks are automatically learnt given images of a monocular camera.
4) The entire navigation is based on sensor data acquired using 3D time-of-flight (ToF) cameras (rather than 2D laser range finders)

In order to ensure an execution quality of storage orders comparable to those of laser-based systems, we divide the drive requests into a (global) approach of a goal pose with respect to the map and an approach of a goal pose within a local coordinate frame. The restricted field-of-view and operating range of the utilized ToF camera entails a number of modifications of common navigation stacks for mobile robots. We will provide insight details and application experiences in the following.

[1]All authors are with the Institute of Computer Engineering, University of Luebeck, Germany.

II. System Overview

Our AGV is equipped with two Basler tof640-20gm cameras with one being mounted in the driving direction and one within the fork. The latter is used for relative pose measurements of pallets aimed to be picked up by the AGV. The following sections provide an overview of the most important components of our navigation system.

A. Data Preprocessing

A point cloud is constructed from the input depth image of our ToF camera. Based on this we generate a 2D scan consisting of the minimal range value returned for each of the 480 columns within a height of $2m$. The limitation to 2D is well-suited for our application as the navigation of the AGV is restricted to motions in 2D assuming a sufficiently planar ground.

B. Mapping

Our system makes use of graph-based simultaneous localization and mapping (SLAM) which has become the de-facto state-of-the-art in mobile robotics. Several algorithms have shown to be capable for the operation at the scale of large warehouses using laser range finders, stereo cameras and to a limited range even using solely monocular cameras. In contrast to laser-based SLAM, we have to cope with a limited field of view, operating range and measurement accuracy, which becomes notable in the sequential scan matching as well as the loop closure detection. This is due to the fact that the limited perception results in a fewer number of overlapping scan points. In contrast to visual SLAM using high-quality 2D camera images, we are solely given low-resolution, noisy intensity images, which do not provide a valuable source for loop closure detection. We therefore implement Geometrical Surface Relations [2] which enable place recognition based on geometrical features extracted from the range scan. In addition to that, we stitch chains of successive scans using scan matching in order to enable a more distinctive place recognition and robust geometric verification thanks to the increased number of point correspondences. While the sequential motion estimation in laser-based SLAM can be conducted using only scan matching, we require a reasonable prior obtained from wheel odometry measurements. This is mainly due to the limited field of view which renders camera-based motion estimation in open space areas impossible. We observed that the availability of an odometry prior significantly contributes to a more robust state estimation. In contrast to visual SLAM based on 2D camera images, our approach is not affected by structure-less parts of the environment (such as plain walls) thanks to the availability of depth data.

C. Localization

Given a map of the environment generated with the above-mentioned SLAM algorithm and a range scan extracted from the input point cloud, we are able to estimate the position of our AGV with respect to the maps coordinate frame. In our work, we leverage the fact that the position accuracy does not have to be equally high at every point in the operation environment. The majority of state-of-the-art navigation systems using laser range finders work in absolute coordinates for approaching pick-up locations, storage points or loading areas. Our approach, in contrast, enables a context-based switch between absolute position estimates w.r.t. the map and relative positioning w.r.t. a target pose. The latter uses a coordinate frame, which is initialized once a drive approach with respect to a local target is started. This allows reducing the precision requirements for our position estimate and enables the use of ToF camera data while the AGV is still able to fulfill all relevant tasks. In particular, we observed that a position accuracy of about $0.5m$ and an orientation accuracy of about 10 suffices for our warehouse application. It has to be ensured that the target object aimed to be approached is visible in the cameras field of view and the system is not confused by ambiguous target measurements. The latter can be observed, for example, when approaching a pallet placed inside a rack. Due to the observation of adjacent pallets, the target recognition module might get confused. Similar to our mapping algorithms, our localization has to cope with the limited field of view of the ToF camera. Thus it is inevitable to incorporate wheel odometry in our state prediction. This provides a beneficial contribution within open-space areas inside warehouses. The increased position uncertainty in such areas can be accepted as our system possesses a fine-positioning system for approaching goal points based on relative pose measurements. Areas that require a higher precision as, for example, within high-level racks, naturally provide more reference contour in close proximity. Our map-based localization uses adaptive Monte-Carlo localization (AMCL) integrated in ROS. In particular it uses a likelihood field sensor model with an increased expected measurement noise which is well-suited for noisy ToF camera data. The relative positioning is accomplished by estimating a relative pose of pallet with a 3D position and a yaw angle using data of the ToF camera placed inside the fork. The pose estimate is frequently updated while the AGV is approaching the target object[2].

III. Experiments

We carried out experiments in a testing warehouse with a size of about $12m\,x\,20m$ consisting of objects which can commonly be found in the intra-logistics such as high-level racks, gates and palleted goods. Our map-based localization using ToF camera data achieves a mean accuracy of about $0.16m$ with a maximum error of about $0.47m$. We further evaluated the overall robustness of our AGV by means of successfully executed storage tasks. As a result we were able to successfully complete 3 out of 4 storage requests. One requests failed due to a position error exceeding $0.4m$.

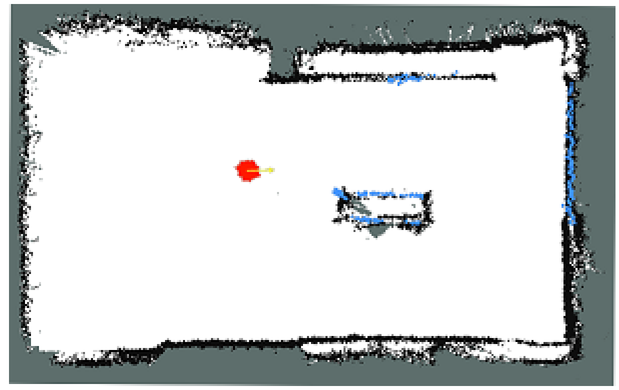

Fig. 1: An example of the localization with respect to a prior map based on AMCL using the ToF camera data is shown. Particles approximating the vehicle's pose (red), the estimated pose (yellow) and the 2D range scan (blue) projected from the ToF camera data are plotted. Note the large field of view of the camera which is not yet common for ToF cameras.

IV. Conclusion

We presented an overview of the navigation system being used in our funded project FTF out-of-the-box. In particular, we showed that two wide-angle ToF cameras are sufficient for the autonomous navigation of a forklift in a warehouse environment. The required modifications for this type of sensors for algorithms being commonly used in mobile robotics were detailed. As an important step we introduced a distinction of approaching points in map coordinates and in a local coordinate frame w.r.t. pose measurements of target objects. This allows to reduce the requirements in regards of position accuracy. Our system was experimentally evaluated in a typical warehouse environment. We hope to motivate the use of ToF cameras for autonomous warehouse navigation since these enable a better object recognition for approaching target objects while simultaneously reducing the overall costs of a navigation systems. In particular, this renders autonomous navigation possible for small to medium sized AGCs and AGVs for intra-logistics applications.

References

[1] L. Overmeyer and F. Podszus and L. Dohrmann, *Multimodal speech and gesture control of AGVs, including EEG-based measurements of cognitive workload*, CIRP Annals - Manufacturing Technology , Vol. 65, pp. 425-428, 2016.

[2] M. Himstedt and E. Maehle, *Geometry matters: Place recognition in 2D range scans using Geometrical Surface Relations*, European Conference on Mobile Robots (ECMR), 2015

[2]Since this component is part of an industry partner's deliverable, we are unable to provide more details at this stage.