Towards an Autonomous Vision-based Inventory Drone

Michael Maurer*, Jesus Pestana*, Friedrich Fraundorfer* and Horst Bischof*

***Abstract*— Logistics is a time and cost intensive task in industry. Great expense is done in the purpose of inventory. To reduce the effort we present an inventory drone prototype developed to demonstrate the feasibility of vision-based navigation inside industrial warehouse facilities with highly repetitive visual structure. Model-based visual localization against a precomputed map of the environment is used to estimate and compensate the drift of a commercially available odometry sensor and to achieve cm-accurate navigation. Our experiments showcase the capabilities of our prototype for vision-based precise indoor navigation and autonomous inventory. All computations are performed on-board the drone.**

I. INTRODUCTION

Nowadays there are still challenges limiting the usage of drones for indoor industrial applications, for instance: the lack of global localization, electromagnetic interference, constrained space and repetitive structure in industrial environments. Thus accurate and global indoor localization and navigation is an active research topic.

In this work, we present a on-board solution for vision-based localization in global coordinates including odometry drift compensation that achieves cm-accurate navigation. In contrast to Beul et al. [1], where a LIDAR sensor is utilized, we focus on a vision-based solution for the localization and navigation problem and on performing the inventory identification using barcodes, rather than RFID tags.

II. METHODOLOGY

The core idea of our proposed solution to achieve precise navigation indoors is to fuse model-based visual localization estimates and the odometry provided by an on-board odometry sensor. As the reference model, depicted in blue in Fig. 1, is metric and has a defined coordinate system, we additionally gain the benefits of localizing in global coordinates, as it is the case outdoors when relying on a Global Navigation Satellite System (GNSS).

To achieve the localization we exploit only the images of a single low resolution camera of the odometry sensor. In preprocessing, the reference model is generated from previously acquired images by using a Structure from Motion approach [2], [3]. Then, during operation the acquired images are localized against the reference model (see Fig. 1). SIFT features are extracted and matched against features of the model. In order to speed up the localization we propagate the current pose according to the odometry estimates and reduce the complexity of matching by only considering nearby reference cameras of the model. Further speed-ups are achieved by tree-based feature matching and GPU-based feature extraction. The matched features implicitly lead to 2D-3D correspondences and the current camera pose is determined using a robust version of the 3-point algorithm.

*The authors are with the Institute of Computer Graphics and Vision, Graz University of Technology, 8010 Graz, Austria [maurer, pestana, fraundorfer, bischof]@icg.tugraz.at

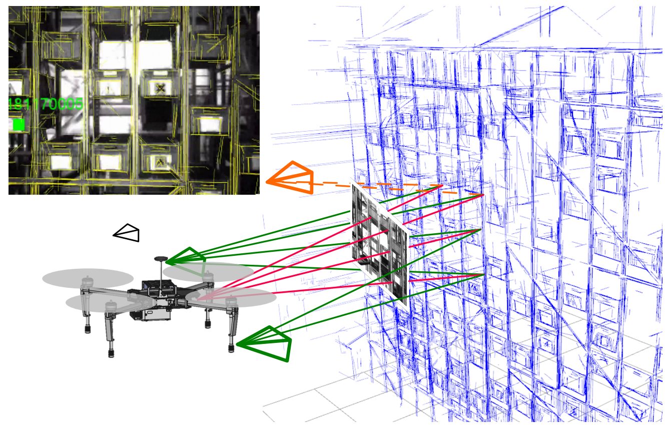

Fig. 1. Principle of vision based real-time localization. The reference model, depicted in blue, is used to localize in real-time. Features of the current image are matched to features of the model's reference images. The feature matching is limited to reference images in the surroundings of the pose hypothesis provided by propagating the odometry starting at the last estimated pose. This results in 2D-3D correspondences that are used to determine the current pose. Reference cameras in the vicinity that provide correspondences are depicted in green, cameras that would provide correspondences but are not close by are colored orange and cameras without any matching features are black. Top-left: Camera image with reprojected reference model and last detected barcode position.

In order to be able to exploit the odometry position and speed estimates to control the drone, we robustly estimate the drift with regards to the reference model of the odometry provided by the on-board sensor. The odometry drift estimation is performed in real-time at very low computational cost. The drift is estimated through a mean calculation over the latest 6 seconds of time-synchronized camera poses and corresponding odometry data. Outliers from the model-based visual localization are filtered using RANSAC. The proposed method enables the compensation of the odometry sensor drift and achieves cm-precise localization.

The controller uses the drift-corrected odometry position and speed estimates as feedback and the acceleration, speed and position of the reference trajectory as feedforward and control setpoints. The reference trajectories are calculated by interpolating a smoothed third order spline over the sequence of desired waypoints.

The barcodes are decoded using the ZXing open-source library[1].

[1]https://github.com/glassechidna/zxing-cpp

III. EXPERIMENTS AND RESULTS

In this section we demonstrate the capabilities of our proposed system by showing the effectiveness of our vision-based drift compensation and summarize our results of autonomous inventory experimental flights.

As development platform we are using a DJI M-100 equipped with a DJI Guidance sensing system [4] for visual odometry that replaces GNSS in indoor scenarios and a Matrix Vision BlueFox3-M1100g (10 Mpx) with a 50 mm focal length lens for barcode reading. The on-board processing is performed on the DJI Manifold (NVIDIA Jetson TK1).

A. Drift Compensation

In indoor scenarios on-board odometry sensors are effected by drift and, if the magnetic compass is fused, by magnetic interferences, which are typical in industrial environments. These deflections can be seen in Fig. 2 in the odometry delivered by the sensor (red curve). We measured a heading drift of 12 deg, which can definitely lead to collisions over long term flight. Thus we exploit our model-based visual localization and robustly fuse the result with the odometry to obtain drift free localization and speed estimates, which enable flying at a constant distance, in Y-direction, to the storage racks (see green curve in Fig 2) with an accurate trajectory tracking error of $\sigma = 0.81$ cm at 0.125 m/s and of $\sigma = 1.30$ cm at 0.55 m/s navigation speed.

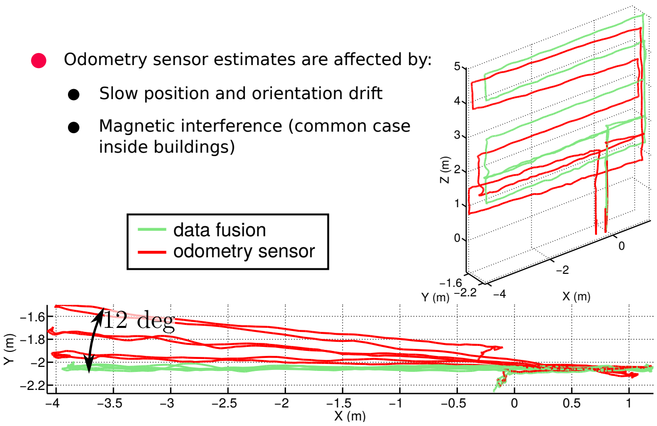

Fig. 2. Odometry delivered by the odometry sensor compared to our drift compensated estimates. Top-right: Odometry drift accumulates particularly in some events such as take-off and landing. Bottom: The magnetic interference causes a slow drift on the heading over long periods of time.

B. Results of the Autonomous Inventory

We show the accuracy of our navigation by the success of the system on reading inventory barcodes of 110 mm length. Thus, we define the constraints to successfully decode a barcode. The inspection camera acquires images of vertical barcodes at a working distance of 1.8 m with a depth of field of 125 mm. At this distance the image spans vertically 225.4 mm and has a resolution of 15.1 px/mm. As we sweep horizontally we only have to consider the field of view in vertical direction. Based on these numbers we can derive the accuracy constraints for the navigation:

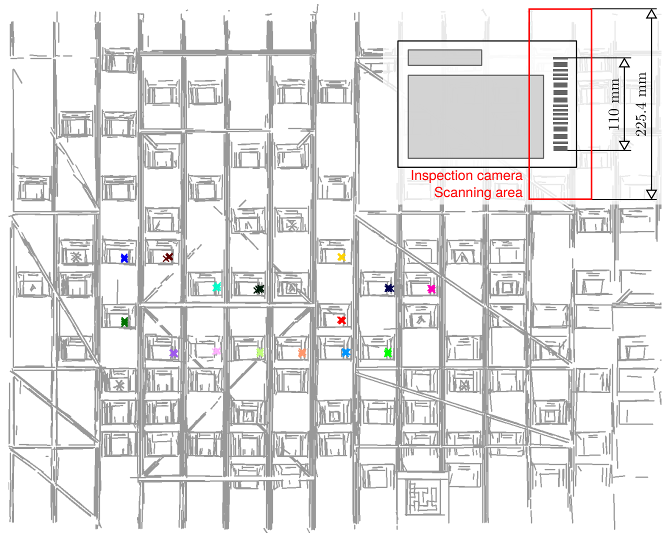

Fig. 3. Positions of detected barcodes on reference model: The colored markers represent the estimated positions of the decoded barcodes for the 5 performed flights. Different colors represent different barcodes. Only the first successful detection of each barcode is considered. Top-right: Field of view of inspection camera and barcode reading constraints.

- vertical tolerance: ± 57.7 mm
- distance tolerance: ± 62.5 mm
- tilt tolerance: ± 1.84 deg

We distributed 15 inventory tags with barcodes on the flight area and performed 5 flights[2]. The inventory drone prototype was able to decode 73 out of 75 barcodes successfully, which shows the precision and repeatability of our navigation. The achieved positioning accuracy of the barcodes in 3D in these flights has a standard deviation of 1.92 cm. Fig 3 shows the individual positions of each barcode clustered by color.

IV. CONCLUSIONS

We presented a solution to achieve vision-based localization in global coordinates with odometry drift compensation using a commercially available odometry sensor. In flight experiments we demonstrated the capabilities of our approach and achieved cm-accurate localization, navigation and position estimation of the inventory barcodes. All computations are performed on-board on our inventory drone prototype.

REFERENCES

[1] M. Beul, D. Droeschel, M. Nieuwenhuisen, J. Quenzel, S. Houben, and S. Behnke, "Fast Autonomous Flight in Warehouses for Inventory Applications," *IEEE Robotics and Automation Letters (RA-L)*, Oct 2018.

[2] M. Rumpler, A. Tscharf, C. Mostegel, S. Daftry, C. Hoppe, R. Prettenthaler, F. Fraundorfer, G. Mayer, and H. Bischof, "Evaluations on multi-scale camera networks for precise and geo-accurate reconstructions from aerial and terrestrial images with user guidance," *Computer Vision and Image Understanding (CVIU)*, vol. 157, pp. 255 – 273, 2016.

[3] M. Hofer, M. Maurer, and H. Bischof, "Efficient 3D scene abstraction using line segments," *Computer Vision and Image Understanding (CVIU)*, vol. 157, pp. 167 – 178, 2017.

[4] G. Zhou, L. Fang, K. Tang, H. Zhang, K. Wang, and K. Yang, "Guidance: A Visual Sensing Platform For Robotic Applications," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2015.

[2] Video of one experiment: https://files.icg.tugraz.at/f/17869e064fc44326a6e0/