Experimental Analysis of Measurements Fusion for Pose Estimation Using PMD Sensor

Ksenia Klionovska$^{(\boxtimes)}$, Heike Benninghoff, Eicke-Alexander Risse, and Felix Huber

German Space Operations Center, Muenchener Str. 20, 82234 Wessling, Germany
{ksenia.klionovska,heike.benninghoff
eicke-alexander.risse,felix.huber}@dlr.de

Abstract. This article presents the experimental investigation of the fusion concept of two relative position and orientation (pose) estimates of the rotating target using single Photonic Mixer Device (PMD) sensor for the frame-to-frame tracking. For each frame PMD depth sensor provides co-registered depth and amplitude images of the scene. We propose to use two different pose estimation techniques for each of the data channel with a further fusion of its measured state vectors. The fusion architecture of the state measurements is based on a low complexity weighted average algorithm. The weights for the fusion operator are calculated experimentally with the real data from PMD sensor mounted on DLR's European Proximity Operations Simulator. The fused state vector obtained with experimental results outperforms the accuracy of the two incoming pose measurements. This allows us to ensure robust pose estimation of a rotating target for the whole tracking.

Keywords: PMD sensor · Data fusion · Pose estimation · Optical navigation

1 Introduction

Computer vision refers to the discipline where the combination of the theory and technology of the artificial systems aims to extract, analyze and understand necessary information from digital image or video. Computer vision is used in different application fields, e.g. in medical computer vision or medical image processing, in industry, for military purposes, for the autonomous land based and aerial vehicles [1]. Currently the scope of our work is directed to computer vision in space applications, namely, visual pose estimation of the target spacecraft during the rendezvous phase for On-Orbit Servicing (OOS) missions. Main duties of OOS mission include, e.g. refuel, repairs/upgrades of some parts of satellite and also deorbiting of no more usable spacecraft. The problem can be addressed as follow: the chaser spacecraft should autonomously navigate to the target by estimating relative position and orientation of the target spacecraft using visual sensor and extracting information from an image or sequence of images.

Different types of visual sensors have been considered and tested for the pose estimation during the rendezvous phase. These are LiDAR [2], monocular and stereo cameras [3] and also time-of-flight sensors based on Photonic Mixer Device (PMD) [4, 5] technology. Since the PMD camera has never been used in space so far, we continue investigating it as one possible candidate for visual navigation in space. In this paper we provide an experimental research on the fusion of the state vectors estimated using PMD sensor and suitable algorithms.

Having a look into a recent state-of-the-art techniques for the sensor data fusion in other different areas of computer vision, one can mention the work of Schramm et al. [6], where the authors present an approach to fuse the data from stereo, depth and thermal cameras for robust self-localization. The resulting position is obtained through the Extenden Kalman Filter (EKF). The Kim et al. [7] show how to fuse the radar and visual images for the advanced driver assistance system via extrinsic calibration process. Deilamsalehy et al. [8] propose to fuse data from LiDAR, camera and Inertial Measurement Unit (IMU) using EKF for pose estimation. Instead of using mentioned EKF or any other filter for the data fusion, we create a distributed system with a weighted average algorithm, where one state vector is calculated with a depth image, whereas the other one independently with an amplitude image. The weighted average approach for the state vector fusion is simple to implement, and moreover, its suitable for any application, since it doesn't require to know a system dynamics.

All experiments presented in this paper were tested in the European Proximity Operations Simulator (EPOS 2.0) [9] at the German Aerospace Center (DLR), which allows real-time simulation of close range proximity operations under realistic space illumination conditions.

2 Visual Navigation with PMD Sensor

2.1 Problem Statement

The problem addressed in this paper is to accurately estimate position and orientation of the target spacecraft with measurements from a time-of-light PMD sensor. In this work we provide experiments with DLR-Argos 3D-P320 Camera prototype (Fig. 1, left up) released by Bluetechnix company with technical characteristic presented in the Table 1.

The depth measurement principle of the PMD technologies is based on the calculation of the phase shift between the emitted NIR signal by the LED's of camera and reflected signal from the target. Co-registered amplitude information of the reflected signal is calculated simultaneously. An example of depth and amplitude images acquired by the DLR-Argos 3D-P320 Camera in DLR's EPOS is presented in Fig. 1 (left and right down images). The images illustrate target mockup (see Fig. 1, right up) used for the further test simulation for this paper. In the next section we provide experimentally justified fusion technique of two pose estimates from different sources in order to get one accurate pose for every frame.
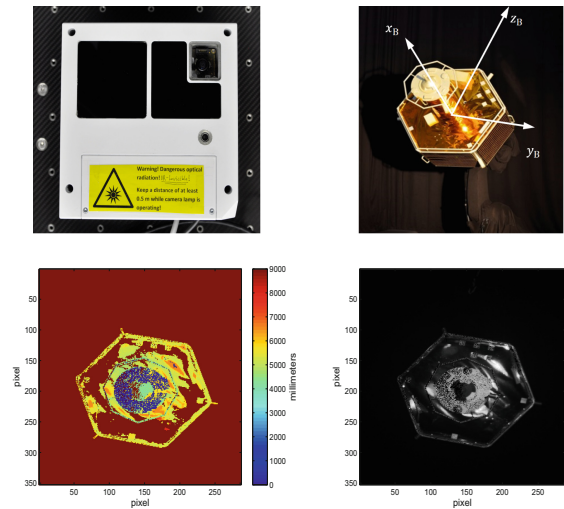
Fig. 1. Left up: DLR-Argos 3D-P320 Camera. Right up: The mockup in EPOS laboratory. Left down: Depth image. Right down: Amplitude image.

Table 1. Technical data of the PMD sensor inside the DLR-Argos 3D-P320 camera.

Field of view	28.91×23.45 deg
Resolution of the chip	352×287 pixels
Integration time	24 ms
Frames per second	45
Modulated frequencies	5.004 MHz, 7.5 MHz, 10.007 MHz, 15 MHz, 20.013 MHz, 25.016 MHz, 30 MHz
Mass/Power consumption	2 kg/$<$25.5 W

2.2 Pose Estimation Techniques

Two completely different tracking methods are suggested for model-based pose estimation techniques of the target spacecraft. Model-based techniques refer to the knowledge of the 3D model of the target object throughout the entire tracking period. The modified version of state-of-the-art Iterative Closest Point (ICP) [10] algorithm with "reverse-calibration" method [11] for the nearest-neighbor search is proposed for the estimation of the state vector with depth images. Pose estimation technique for the amplitude images based on finding feature correspondences between 3D known model and detected features in the 2D gray scaled image. Throughout variety of the solvers here we propose to take a Gauss-Newton solver [12] based on a least square minimization problem in order

to estimate position of the target related to the camera frame. Please, refer to the work of Klionovska et al. [5] for the detailed description of the pose estimation technique for the depth images and for the feature identification from amplitude images.

2.3 Fusion of Measurements

Due to the redundant information from the PMD sensor (depth and amplitude channels) it is a good chance to increase the reliability of the system and enhance the accuracy of the calculated state vector during the approach by fusion of two estimated measurements. Moreover, with the redundant state information the tracking of the target can still be provided even when one of the pose estimation techniques gives incorrect information. One of the simplest ways for the combination of measurements is to take a weighted average [13] of the pose vectors which is obtained after two different pose estimation techniques. The simple arithmetic mean of all measurements does not perform enough, since one measurement can be more reliable than other [14]. Taking in account this fact, it is better to assign more importance and greater weight to an observation y_i from one output channel that is more reliable, whereas a less accurate observation from the other output channel will receive minor weights. The weighted average for the fused estimate of n different measurements y_i with non-negative weights ω_i looks as

$$y_{fused} = \frac{\sum_{i=1}^{n} \omega_i y_i}{\sum_{i=1}^{n} \omega_i}. \tag{1}$$

We can simplify an Eq. 1 when the weights are normalized and sum up to 1:

$$y_{fused} = \sum_{i=1}^{n} \omega_i' y_i, \sum_{i=1}^{n} \omega_i' = 1. \tag{2}$$

From the mathematical point of view the weights ω_i for every single member of the pose vector can be assigned as estimated variance of the measurement error σ_i^2 occurred during pose estimation with one of the suggested methods

$$\omega_i = \frac{1}{\sigma_i^2}. \tag{3}$$

In the work of Elmenreich [15] the author shows that the variance of the fusion result y_{fused} is minimized and always smaller than the input variances

$$\sigma_{fused}^2 = \sum_{i=1}^{n} \omega_i^2 \sigma_i^2 = \sum_{i=1}^{n} \frac{1}{\sigma_i^2}. \tag{4}$$

3 Experimental Scenarios and Performance Analysis

In order to find correct weights for fusion of both estimates, firstly we propose to run both algorithms separately. The offline test presents a straight frontal approach scenario, which starts at approximately distance 8 m between chaser and the target. The termination point is situated at the distance a bit less than 5 m. These both distances are chosen due to the characteristics of the DLR-Argos 3D-P320 sensor. Namely, the starting point of the simulation is chosen with relation to the resolution of the current PMD sensor. Since the resolution of the PMD sensor is small in compare with existent CCD sensors, the features of the imaged object become not to be clearly observable and it leads to the big errors in pose estimation. Moreover, the current illumination unit of the camera is suitable for the close range simulations (<10 m) and not for the long one. The final point is limited because of the field of view of the current camera. When the distance is less than 5 m the whole target is no more observable and pose estimation is not possible. In this test the target (Fig. 1, right up) is rotating around its principal axis of inertia at a rate of 2 deg/s. Overall the data set consists of 170 images. According to the number of image, the distance range is configured as following: from image 1 to 67 corresponds to the approach from 8 to 7 m; from image 68 to 118 is a range from 7 to 6 m (e.g. Fig. 2, images at the first row); and starting from image 119 to 170 the distance decreases from 6 to 4.9 m (e.g. Fig. 2, images at the second row). The ground truth for every logged image is provided by EPOS. The experimental scenarios, which we consider in this paper, are follows. Test scenario 1 presents frame-to-frame pose estimation technique using depth image and ICP with reverse calibration technique. Test scenario 2 concludes the result of the pose estimation algorithm with aforementioned image processing of amplitude images. Test scenario 3 shows the results of the fusion technique with calculated weights. What is special here is that in the Test scenario 3 the fusion technique is applied for the translational part, whereas the rotation is completely taken from the pose vector estimated with the amplitude image. This is so, because the algorithm with the amplitude images is less sensitive for the estimation of the orientation and usually provides better results. The camera coordinate frame is used in order to evaluate the results, where Z-axis is taken along the optical axis of the camera.

3.1 Test Cases 1 and 2: Pose Estimation Using Depth and Amplitude Images Separately

We run separately the algorithms for the depth images and amplitude images using provided dataset. In Fig. 3 we present the plots of the errors for rotation and translation components of the estimated pose for every frame in the Test case 1 (left up and down images) for the Test case 2 (right up and down images). The mean errors for the both cases are presented in the Table 2. From the results depicted in Fig. 3 and collected in Table 2 after offline simulation of the proposed estimation techniques explained in the Sect. 2.2, one can observe that estimation

of the rotational components with the 2D technique using amplitude image dominates over the 3D pose estimation pipeline. It is due to the fact that the errors in the angles calculated with 3D pose estimation algorithm have a tendency to accumulate. This is caused by the nature of the algorithm - since the previous estimate for a new frame strong diverges from the real one, the follow calculated orientation (sometimes also position vector) within a next new frame has also big measurement errors. However, position of the target during the tracking was defined more accurate using the depth images, especially the distance component (position along Z axis).
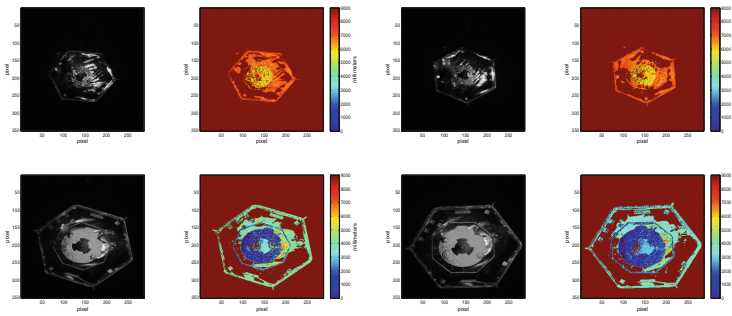
Fig. 2. Depth and amplitude images within distance 7 to 6 m (first row) and within distance 6 to 5 m (second row).
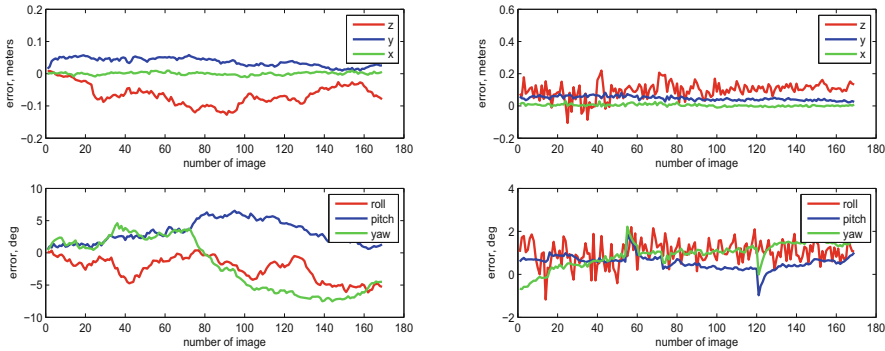
Fig. 3. Translation and rotation errors for the Test cases 1 (left up and down) and 2 (right up and down).

In order to apply the fusion technique described in the Sect. 2.3, there is a prerequisite to define weights. According to the Eq. 3, it is necessary to define the variances. Table 2 presents the result of the standard deviations for three rotation angles and for Z, Y, X components of the position vector.

Table 2. Standard deviations and mean errors.

σ	Test case 1	Test case 2	μ	Test case 1	Test case 2	Test case 3
σ_{roll}, deg	1.75	0.47	μ_{roll}, deg	2.41	1.04	0.96
σ_{pitch}, deg	1.74	0.27	μ_{pitch}, deg	3.19	0.57	0.62
σ_{yaw}, deg	2.18	0.43	μ_{yaw}, deg	3.91	0.98	1.10
σ_z, m	0.0306	0.0415	μ_z, m	0.0654	0.1004	0.0392
σ_y, m	0.0129	0.0129	μ_y, m	0.0366	0.0472	0.0405
σ_x, m	0.0024	0.0058	μ_x, m	0.0034	0.0065	0.0038

Table 3. Weights for the translation components.

Weight	Depth estimate	Amplitude estimate
ω_z	0.65	0.35
ω_y	0.54	0.46
ω_x	0.85	0.15

3.2   Test Cases 3: Fusion of Pose Vectors with Weights

Taking into account revealed tendency after test cases 1 and 2, we propose to make simulations of the pose estimation during the approach for the same dataset and by using the weighted average technique for the translation components. We apply weights for the Z, Y and X coordinates presented in the Table 3. In Fig. 4 we plot the results of the angular and position measurement errors after conducted Test case 3.
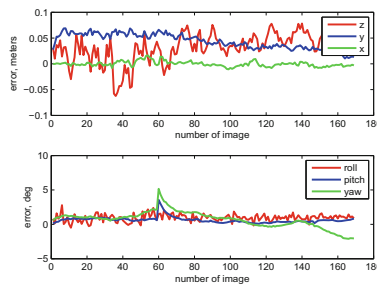
Fig. 4. Translation and rotation errors for the Test cases 3.

In fact, as we expected, the fused technique with its measurements errors presented in Fig. 4 overcomes the drawbacks of both pose estimation techniques. The mean errors for Test case 3 are shown in the last column of Table 2. It means that the attitude of the target has almost the same mean errors as in the Test

case 2, whereas the mean errors for the position are more similar with the Test case 1. The peaks for the pitch and yaw angles presented in Fig. 4 don't corrupt or abrupt the tracking process, allowing reliably continue pose estimation of the target.

4   Discussion

With this paper we proved experimentally that the fusion technique guarantee accurate calculation of the position and orientation of the target using PMD sensor. Here for every frame the weighted average method fuses two estimates, which were calculated with depth and amplitude images independently. By different test simulations we have shown the main advantage of the fusion - decrease of the measured errors for the attitude and for the position of the target. Moreover, having a fused estimate we ensure stable frame-to-frame tracking during the approach. The proposed concept of the data fusion for pose estimation and tracking based on PMD sensor technologies can be used not only for space applications. For example, independent redundant information from one PMD visual sensor can be used in the field of (semi)autonomous driving and also driver assistance systems. This approach together with PMD sensor helps to reduce the number of visual sensors, but at the same time ensures increase of the accuracy and reliability of the visual system.

References

1. Examples of Applications for Computer Vision. http://www.idc-online.com/technical_references/pdfs/electronic_engineering/Examples_of_applications_for_computer_vision.pdf
2. Opromolla, R., Fasano, G., Rufino, G., Grassi, M.: Uncooperative pose estimation with a LIDAR-based system. Acta Astronautica **110**, 287–297 (2015)
3. Oumer, N.W.: Visual Tracking and Motion Estimation for an On-Orbit Servicing of a Satellite (2016)
4. Ringbeck, T., Hagebeuker, B.: A 3D Time of Flight Camera for object detection. Optical 3D Measurement Technologies (2007)
5. Klionovska, K., Ventura, J., Benninghoff, H., Huber, F.: Close Range Tracking of an Uncooperative Target in a Sequence of Photonic Mixer Device (PMD) Images (2018). https://doi.org/10.3390/robotics7010005
6. Schramm, S., Rangel, J., Kroll, A.: Data fusion for 3D thermal imaging using depth and stereo camera for robust self-localization. In: IEEE Sensors Applications Symposium (SAS) (2018)
7. Kim, J., Han, D.S., Senouci, B.: Radar and vision sensor fusion for object detection in autonomous vehicle surroundings. In: Tenth International Conference on Ubiquitous and Future Networks (2018)
8. Deilamsalehy, H., Havens, T.C.: Sensor fused three-dimensional localization using IMU, camera and LiDAR. IEEE Sensors (2016)
9. Benninghoff, H., Rems, F., Risse, E.-A., Mietner, C.: European Proximity Operations Simulator 2.0. (EPOS)- a robotic-based rendezvous and docking simulator. J. Large Scale Res. Facil. **3**, A107 (2017)

10. Mckay, N., Besl, P.: A method for registration of 3d shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**, 239–256 (1992)
11. Blais, G., Levine, M.: Registering multiview range data to create 3D computer objects. IEEE Trans. Pattern Anal. Mach. Intell. **17**, 820–824 (1995)
12. Nocedal, J., Wright, S.: Numerical Optimization. Springer, New York (2006). https://doi.org/10.1007/978-0-387-40065-5
13. Grossman, J., Grossman, M., Katz, R.: The First Systems of Weighted Differential and Integral Calculus. Archimedes Foundation, Rockport (1980)
14. Schrgendorfer, A., Elmenreich, W.: Extended confidence-weighted averaging in sensor fusion. In: Proceedings of the Junior Scientist Conference (2006)
15. Elmenreich, W.: Fusion of continuous-valued sensor measurements using confidence-weighted averaging. J. Vib. Control. **13**, 1303–1312 (2007)