

Precise Positioning of Robots with Fusion of GNSS, INS, Odometry, Barometer, Local Positioning System and Visual Localization

Patrick Henkel ^{*,**}, Andreas Sperl ^{*}, Ulrich Mittmann ^{*}, Robert Bensch ^{*}, Paul Färber ^{*} and Christoph Günther ^{*,**,***}

^{*}*Technische Universität München (TUM), Munich, Germany*

^{**}*ANavS GmbH - Advanced Navigation Solutions, Munich, Germany*

^{***}*German Aerospace Center (DLR), Oberpfaffenhofen, Germany*

BIOGRAPHIES

Patrick Henkel received his B.Sc. and M.Sc. in Electrical and Computer Engineering and his PhD from the *Technische Universität München*, Munich, Germany. In 2010, he graduated with a PhD thesis on reliable carrier phase positioning (summa cum laude) and is now working towards his habilitation in the field of precise point positioning. He visited the Mathematical Geodesy and Positioning group at TU Delft in 2007, the GPS Lab at Stanford University in 2008 and 2010, and the GNSS Research Centre at Curtin University in 2013. Dr. Henkel received the Pierre Contensou Gold Medal in 2007, the 1st prize in Bavaria at the European Satellite Navigation Competition in 2010, and the Vodafone Award for his dissertation in 2011. He is one of the founders and currently also the managing director of ANavS GmbH - Advanced Navigation Solutions.

Andreas Sperl studied Electrical and Computer Engineering at the *Technische Hochschule Deggendorf* and the *Technische Universität München*, Munich, Germany. His master thesis was entitled *Joint RTK and Attitude determination* and evaluated with the grade "with distinction". Mr. Sperl joined ANavS GmbH as software and hardware developer in 2015, and is currently a shareholder and head of development at ANavS GmbH.

Ulrich Mittmann completed his apprenticeship in energy electronics at Krauss Maffei Group GmbH. He studied at the *Hochschule für angewandte Wissenschaften München* and received his B. Sc. and M. Sc. in electrical engineering from the *Technische Universität München*. In his Master's thesis he investigated the GLONASS ambiguity resolution for RTK positioning. Since 2016, he is working for ANavS GmbH as algorithm and software developer on Multi-Sensor RTK positioning.

Robert Bensch received a Bachelor in information technology, and the Master and PhD degree in applied computer science from the *Albert-Ludwigs-Universität Freiburg*, Freiburg, Germany. In his research he focused on biomedical image analysis under the guidance of Prof. Burkhard, Prof. Brox and Prof. Ronneberger as a member of their computer vision group and image analysis lab. He received funding from BIOSSE, Centre for Biological Signalling Studies (The Excellence Initiative of the German Federal and State Governments), and contributed to the study of biological signalling processes. In 2017, he graduated with his PhD thesis on motion pattern analysis for biomedical applications. He joined ANavS GmbH as a software architect and is currently heading the computer vision team.

Paul Färber received his B.Sc. and M.Sc. in Electrical and Computer Engineering from the *Technische Universität München*. He graduated in 2018 with a masters thesis on the integration of visual odometry into the multi-sensor fusion. He is currently working in the area of precise positioning in different environments for safe navigation at ANavS GmbH.

Christoph Günther studied theoretical physics at the Swiss Federal Institute of Technology in Zurich. He received his diploma in 1979 and completed his PhD in 1984. He worked on communication and information theory at Brown Boveri and Ascom Tech. From 1995, he led the development of mobile phones for GSM and later dual mode GSM/ satellite phones at Ascom. In 1999, he became head of the research department of Ericsson in Nuremberg. Since 2003, he is the director of the Institute of Communication and Navigation at the German Aerospace Center (DLR) and since December 2004, he additionally holds a Chair at the *Technische Universität München*. His research interests are in satellite navigation, communication and signal processing.

ABSTRACT

The autonomous driving of robots requires precise and reliable positioning information with low-cost sensors. In this paper, we propose a tightly coupled sensor fusion of multiple complementary sensors including GNSS-RTK, INS, odometry, Local Positioning System (LPS) and Visual Positioning. The focus of this paper is on the integration of LPS and vision since the coupling of GNSS-RTK, INS and odometry is already state of the art. We include the positions of the LPS anchors and the bearing vectors and distances from the robot's camera towards the patch features as state vectors in our Kalman filter, and show the achievable positioning accuracies.

INTRODUCTION

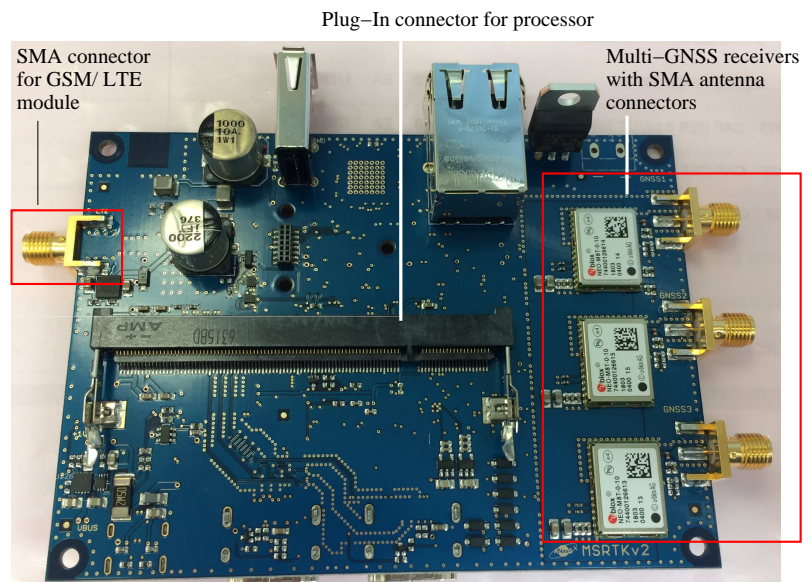
The purpose of this section is two-fold: First, we introduce the complementary properties of 5 positioning sensors: Global Navigation Satellite System (GNSS) receivers, Inertial Measurement Units (IMU), wheel odometry, Local Positioning Systems (LPS), and camera-based Visual Positioning. Second, we present the ANavS Multi-Sensor module [1] that carries multiple low-cost sensors, communication interfaces and a processor for performing the sensor fusion. Tab. 1 lists the 5 positioning sensors and the advantageous/ challenging environments for each sensor.

Tab. 1: Comparison of complementary positioning sensors: description of conditions resulting in high performance and of conditions resulting in poor performance for each individual sensor.

Sensor	Conditions enabling a high positioning accuracy	Conditions resulting in poor positioning accuracy
GNSS	open-sky conditions with at least 4 visible satellites with continuous phase tracking	any area with less than 4 visible satellites with continuous phase tracking (e.g. below trees, bridges or tunnels)
Inertial Measurement Unit (IMU)	any area for a few seconds after initialization	any area <i>after</i> a few seconds after initialization
wheel odometry	any area with paved roads	gravel-ground roads (slippage)
Local Positioning System (LPS)	any area with line of sight to <i>at least</i> 3 anchors	any area with line of sight to <i>less than</i> 3 anchors
Visual positioning	any area with clear textures, e.g. road markings and road signs,	during fog or heavy snowfall, resulting in camera images without textures

In this paper, the new Multi-Sensor Fusion RTK module of ANavS is used for positioning. The module is shown below and offers the following key features [1]:

- 1 to 3 integrated GNSS receivers for RTK positioning and attitude determination
- integrated inertial sensor and barometer for robust positioning
- integrated CAN-bus interface for odometry and CSI-interface for camera
- integrated LTE module for reception of RTK corrections
- integrated processor for Multi-Sensor, Multi-GNSS tightly coupled RTK positioning



LOCAL POSITIONING SYSTEM

In this section, we describe the Local Positioning System (LPS) and its integration into the sensor fusion. There are two types of LPS range measurements: The first type of range measurements refers to the range between a certain anchor (with index k) and the user/ robot (with index u), and is modeled as

$$r_u^k = \|\vec{x}_u - \vec{x}^k\| + \Delta r_{\text{MP}_u}^k + \eta_u^k = (\vec{e}_u^k)^\top (\vec{x}_u - \vec{x}^k) + \Delta r_{\text{MP}_u}^k + \eta_u^k, \quad (1)$$

with the following notations:

\vec{x}_u	user/ robot position
\vec{x}^k	anchor position
$\vec{e}_u^k = \frac{\vec{x}_u - \vec{x}^k}{\ \vec{x}_u - \vec{x}^k\ }$	normalized direction vector between anchor and robot
$\Delta r_{\text{MP}_u}^k$	multipath error of LPS range measurement
η_u^k	noise of LPS range measurement

The second type of range measurements refers to the anchor-to-anchor measurements. The range measurement between anchors k and l is modeled similar to Eq. (1) as

$$r^{kl} = \|\vec{x}^k - \vec{x}^l\| + \Delta r_{\text{MP}_u}^{kl} + \eta_u^{kl} = (\vec{e}^{kl})^\top (\vec{x}^k - \vec{x}^l) + \Delta r_{\text{MP}_u}^{kl} + \eta_u^{kl}. \quad (2)$$

We determine the positions of the robot and all anchors jointly in a Kalman filter. The anchor-to-robot and anchor-to-anchor measurements are stacked in a single column vector as

$$z = \left(r_u^1, \dots, r_u^K, r^{12}, \dots, r^{1K}, \dots, r^{(K-1)K} \right)^\top, \quad (3)$$

which includes $K + \frac{(K-1)K}{2}$ linear independent measurements.

The LPS can not provide a unique solution, i.e. the positions of all anchors and the robot can be shifted by a common, arbitrary vector and be rotated by a common, arbitrary rotation matrix without affecting the range measurements. This leaves 6 degrees of freedom. We use them

- to set the coordinate center of the Local Positioning System (LPS) to the position of the first LPS anchor
- to define the x-axis of the coordinate frame, such that it points from the coordinate center towards the second LPS anchor
- to define the y-axis of the coordinate frame, such that it lies in the plane spanned by the first two anchors and the third anchor
- to define the z-axis of the coordinate frame, such that it complements a right-hand coordinate frame

Fig. 1 shows the LPS coordinate frame based on the positions of three anchors.

The $\{x, y, z\}$ coordinates of the anchor positions in the LPS coordinate frame are noted as

$$\vec{x}^1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \vec{x}^2 = \begin{pmatrix} d_x^2 \\ 0 \\ 0 \end{pmatrix}, \vec{x}^3 = \begin{pmatrix} d_x^3 \\ d_y^3 \\ 0 \end{pmatrix}, \vec{x}^4 = \begin{pmatrix} d_x^4 \\ d_y^4 \\ d_z^4 \end{pmatrix}, \dots, \vec{x}^K = \begin{pmatrix} d_x^K \\ d_y^K \\ d_z^K \end{pmatrix}, \quad (4)$$

where 6 coordinates are 0. The remaining $3(K-2)$ coordinates are unknown and have to be estimated in the Kalman filter.

The state vector comprises the unknown position coordinates of the robot and all anchors, i.e.

$$x = (\vec{x}_u^\top, d_x^2, d_x^3, d_y^3, (\vec{x}^4)^\top, \dots, (\vec{x}^K)^\top)^\top. \quad (5)$$

The anchor positions are assumed to be constant and the robot is assumed to move with a low speed. Thus, the state space model is straight forward, i.e. we assume constant state parameters. The change of the robot's position is accounted for in the process noise.

We use a standard Kalman filter [6] for the Local Positioning System (LPS), and consider the following aspects:

- iterative approach required for state update due to linearization of range measurements
- certain movement required for convergence of positions of anchors and robot

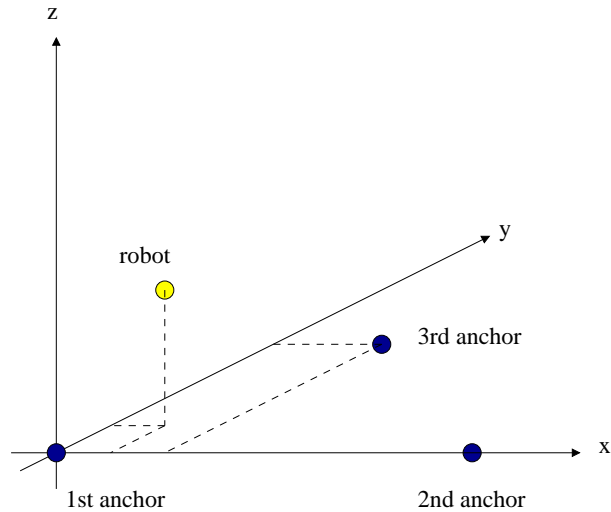


Fig. 1: Local Positioning System (LPS): The right-hand coordinate frame is spanned by the locations of three anchors.

INTEGRATION OF LOCAL POSITIONING SYSTEM INTO SENSOR FUSION

In this section, we briefly describe the sensor fusion of the Local Positioning System (LPS), the Inertial Measurement Unit (IMU), and wheel odometry. We perform a tightly coupled sensor fusion as shown in Fig. 2, i.e. the raw measurements of all sensors are directly used to estimate the state vector comprising the position, velocity, acceleration, attitude angles and angular rates of the robot, the anchor positions, and the IMU and odometry biases. A standard extended Kalman filter is used for the sensor fusion as described by Brown and Hwang in [6].

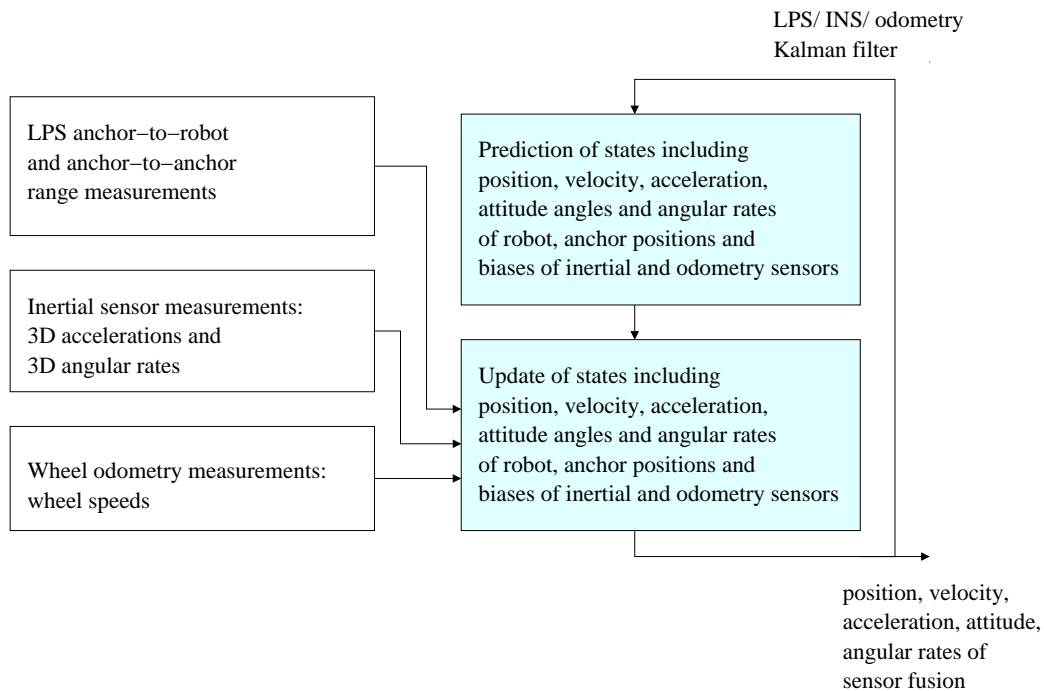


Fig. 2: Sensor Fusion of LPS, INS and Odometry: A Kalman filter is used to predict and update the state vector at every measurement epoch. The state vector includes the position, velocity, acceleration, attitude angles and angular rates of the robot, the anchor positions, and the IMU and odometry biases.

VISUAL-INERTIAL ODOMETRY

In this section, we describe some fundamentals for visual-inertial odometry with a monocular camera.

Projection model and linear warping

In this subsection, we discuss the relationship between the pixel coordinates of a landmark and the bearing vector pointing from the robot to the landmark. We use the Robust Visual-Inertial Odometry (ROVIO) method of Blösch et al. and closely follow their description in [2]-[4]. The pixel coordinates p_n^l of landmark l in frame n are expressed in terms of the camera model π with known intrinsic calibration, and the bearing vector μ_n^l of the landmark:

$$p_n^l = \pi(\mu_n^l). \quad (6)$$

Solving this equation for μ_n^l yields

$$\mu_n^l = \pi^{-1}(p_n^l). \quad (7)$$

The bearing vector is predicted to the next camera frame with a certain process model, i.e.

$$\mu_{n+1}^l = f(\mu_n^l), \quad (8)$$

and then re-projected to pixel coordinates:

$$p_{n+1}^l = \pi(\mu_{n+1}^l) \quad (9)$$

Concatenating the projections of Eq. (7) to (9) relates the pixel coordinates of a certain landmark in two subsequent frames:

$$p_{n+1}^l = \pi(f(\pi^{-1}(p_n^l))) \quad (10)$$

We linearize these projections for the Kalman filter and obtain the following linear warping matrix:

$$D = \frac{\partial p_{n+1}^l}{\partial p_n^l} = \frac{\partial \pi(\mu_{n+1}^l)}{\partial \mu_{n+1}^l} \frac{\partial f(\mu_n^l)}{\partial \mu_n^l} \frac{\partial \pi^{-1}(p_n^l)}{\partial p_n^l}. \quad (11)$$

Photometric error

The photometric error is defined as the pixel-wise intensity difference between a (given multi-level) patch feature and the related image, whereas we have to account for the warping of the patch feature due to a frame (as described in the previous subsection) and eventual changes in illumination between the different frames/ levels:

$$e_{n,j}^l(p^l, P_n, I_n, D) = P_n(p_j^l) - aI_n(p^l s_n^l + Dp_j^l) - b, \quad (12)$$

with the following notations:

P_n	intensity of multi-level patch at frame n
p^l	coordinates of l -th patch feature relative to center of image
p_j^l	coordinates of patch pixel of patch feature relative to center of patch feature
I_n	intensity of image at frame n
s_n^l	scaling factor accounting for downsampling
a	intensity model parameter to account for changes in illumination
b	intensity model parameter to account for changes in illumination

This photometric error is used directly as measurement to update the state vector in our tightly-coupled sensor fusion with a Kalman filter.

INTEGRATION OF VISUAL-INERTIAL ODOMETRY INTO SENSOR FUSION

In this section, we describe the integration of the visual-inertial odometry into the sensor fusion. Fig. 3 shows the architecture for the sensor fusion of GNSS, INS, wheel-odometry and visual-inertial odometry. The LPS measurements are not considered in this section since both GNSS and LPS provide position information.

The visual-inertial odometry uses a Kalman filter that processes the images from a monocular camera and the measurements from an inertial sensor. Our implementation is based on the ROVIO (RObust Visual-Inertial Odometry)-framework of Blösch et al. [2]-[4], that tracks the *bearing vector* and *distance of each patch feature* as state parameter besides the position, velocity, attitude and biases of the inertial sensor. The individual steps of the visual-inertial odometry are highlighted in red. The first step includes the prediction of the state parameters using inertial measurements. Subsequently, the locations of the feature patches are searched in the new camera image around the predicted locations of the feature patches. Finally, the state vector is updated based on the found feature patches.

The obtained position, velocity and attitude estimates serve as measurements for the main Kalman filter, that also uses the GNSS-, INS- and wheel-odometry measurements to update its state vector. The state vector of the main Kalman filter includes the position, velocity, acceleration, attitude angles, angular rates, carrier phase ambiguities, pseudorange multipath errors, and biases of the inertial sensor and wheel odometry. A standard Kalman filter [6] is used for this overall sensor fusion and the respective state prediction and state update steps are highlighted in blue in Fig. 3.

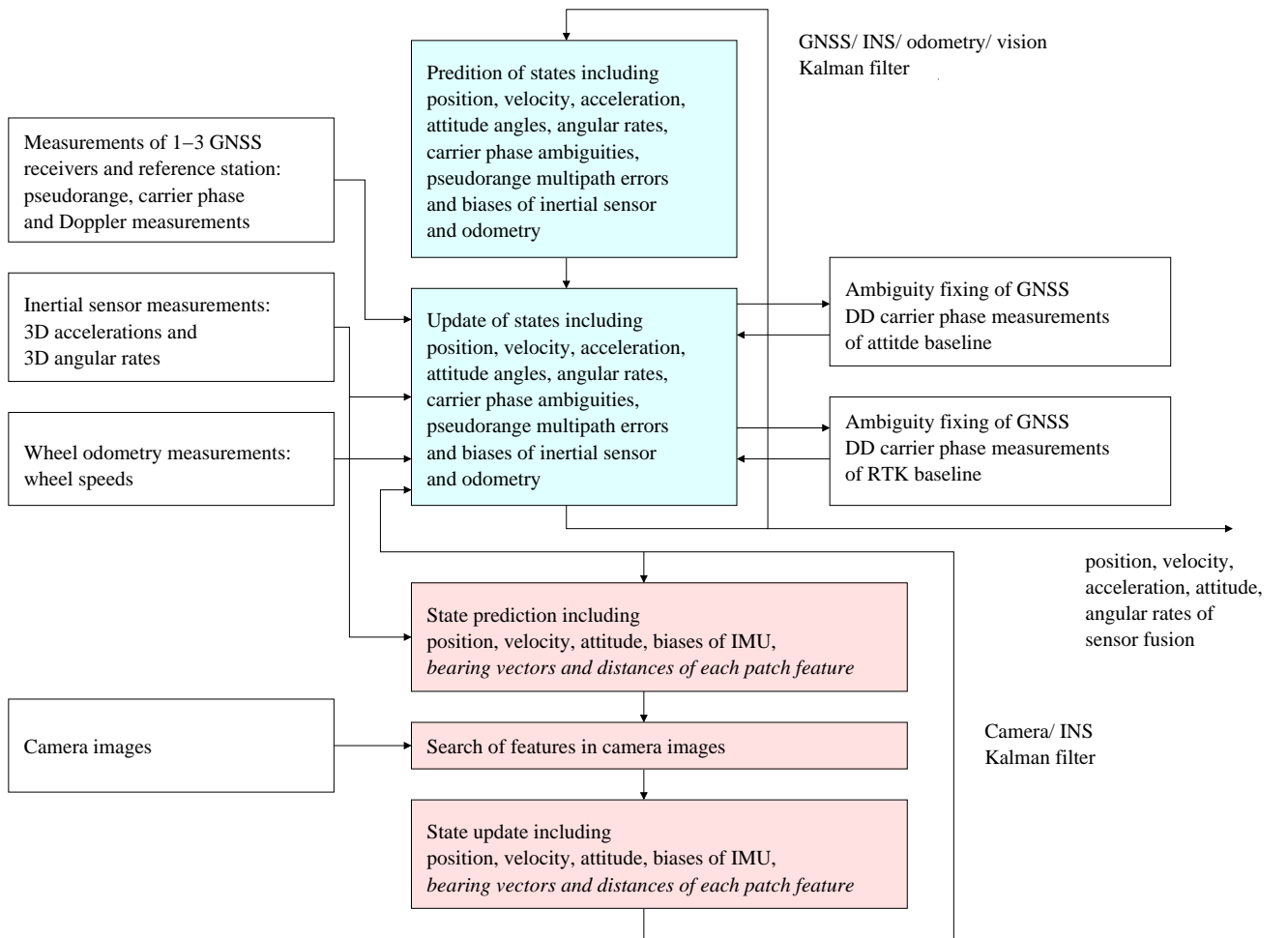


Fig. 3: Architecture for Sensor Fusion of GNSS, INS, wheel-odometry and visual-inertial odometry in Kalman filter.

MEASUREMENT RESULTS

Performance of Local Positioning System

Fig. 4 shows the performance of the Local Positioning System (LPS) with a model train. The positions of three anchors and the robot are jointly estimated. The track of the model train is a closed loop, which enables an analysis of the repeatability of the position solution. The enlarged view provides two insights: First, the point cloud at (3.3 m, 1.95 m) refers to the initial static position, and has a standard deviation of a few centimeters only. Second, the multiple parallel lines refer to different rounds of the model train and indicate a consistent position solution.

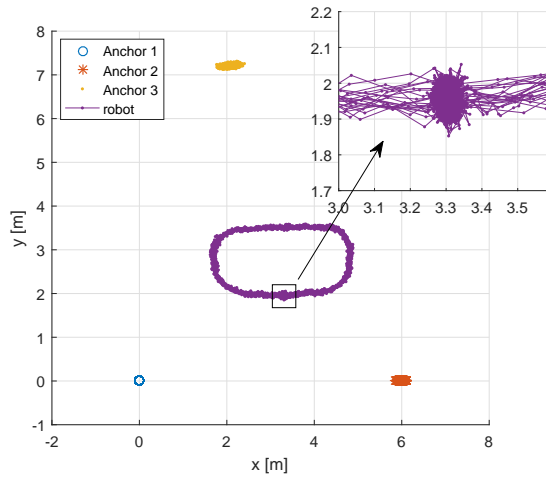


Fig. 4: Performance analysis of Local Positioning System (LPS) with a model train.

Comparison of Local Positioning System and Multi-GNSS Real-Time Kinematic Positioning

Fig. 5 includes a comparison between the ANavS tightly coupled LPS/ INS and the ANavS tightly coupled Multi-GNSS/ INS RTK positioning [7] - [9]. The closed-loop track is installed at a roof-top with open-sky conditions, i.e. both satellite signals and anchor signals are received without obstructions. Both systems are coupled with an IMU and provide consistent solutions with an uncertainty of less than 10 cm for most epochs. The systematic offsets between both positioning solutions around the lower left part and also at the rightmost part of the track are LPS errors that occur if the angle between an LPS antenna plane and the signal path is very small.

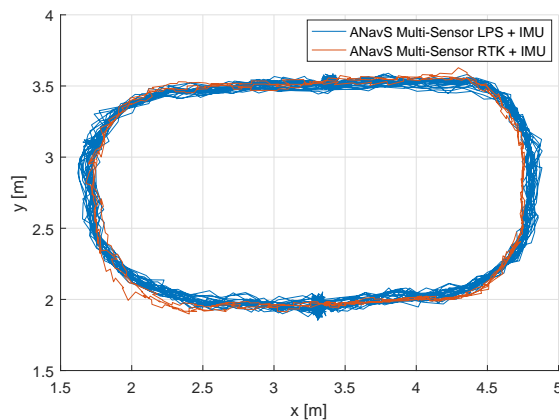


Fig. 5: Comparison between ANavS tightly coupled Multi-GNSS/ INS RTK positioning and ANavS tightly coupled LPS/ INS positioning with model train.

Fusion of Local Positioning System, Wheel Odometry and Inertial Measurements

Fig. 6 shows a comparison between the horizontal position estimates of the ANavS Multi-Sensor Fusion of Local Positioning System, wheel odometry and inertial sensor and the horizontal position estimates of the tachymeter-based reference solution. In principle, both solutions are well aligned for almost all epochs. A slight offset of the LPS/Odo/IMU solution can be observed near the start at (0,0) since the Kalman filter needs some time to converge. The tachymeter solution has occasional gaps due to the lack of a line of sight between tachymeter and robot. Moreover, a temporary reduction of accuracy can be observed for the LPS/ODO/IMU solution in areas where the LPS signals from at least one anchor point were shadowed or blocked, e.g. around (1.5 m, -3.0 m).

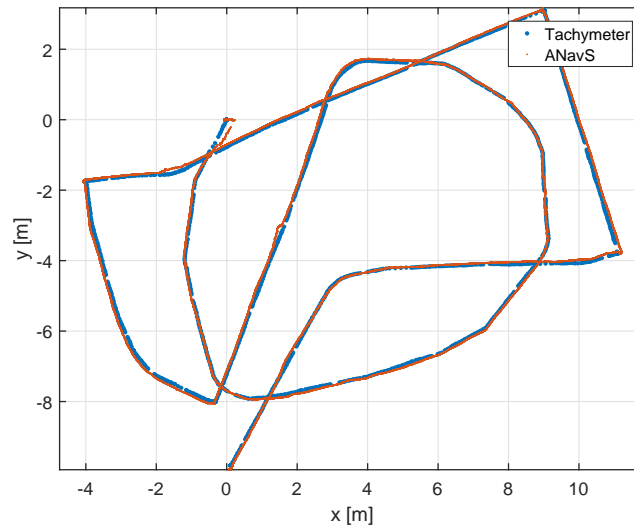


Fig. 6: Comparison of horizontal position estimates between the ANavS Multi-Sensor Fusion of Local Positioning System, wheel odometry and inertial sensor with the tachymeter-based reference solution.

Fig. 7 includes a quantitative assessment of the positioning accuracy of the LPS/ INS/ odometry sensor fusion using a tachymeter as reference. The position offset remains below 15 cm for 95 % of the epochs.

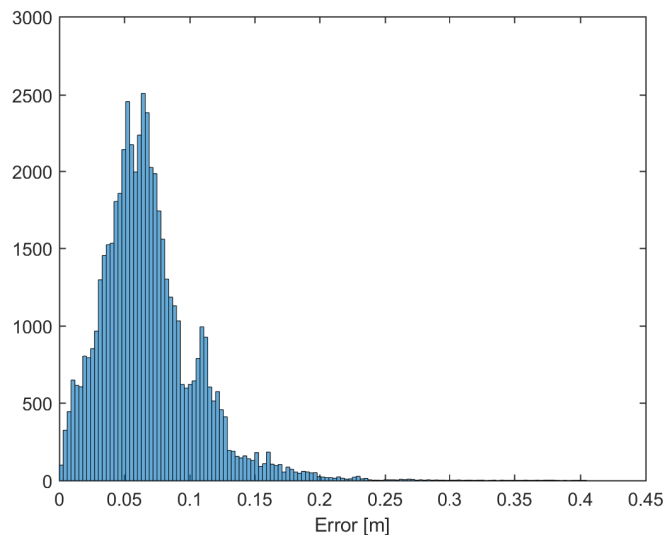


Fig. 7: Histogram of horizontal position deviation between ANavS Multi-Sensor Fusion of Local Positioning System, wheel odometry and inertial sensor and tachymeter-based reference.

Integration of Visual Positioning into Multi-GNSS RTK/ Odometry/ IMU tightly coupled positioning

Fig. 8 shows a comparison of the Multi-GNSS/ wheel odometry/ IMU tightly coupled RTK positioning *with* and *without* integrated visual odometry: The trajectory starts with a rectangular, repetitive pattern at an open field. The initial convergence of the RTK float solution is also shown. The position estimates with and without visual positioning are well-aligned. This indicates the correctness of positioning with and without visual odometry. After the rectangular pattern, the robot drove towards trees and bushes (upper part of trajectory) to test the positioning performance in more challenging conditions. We can observe a certain deviation between the position trajectories with and without visual odometry. The benefit of the visual odometry becomes apparent at the RTK re-fixing after passing the sections with trees and bushes: The position correction is only 20 cm with visual odometry compared to 30 cm without visual odometry. The diagram also shows three highlighted locations. The respective camera images are provided in Fig. 9 and 10.

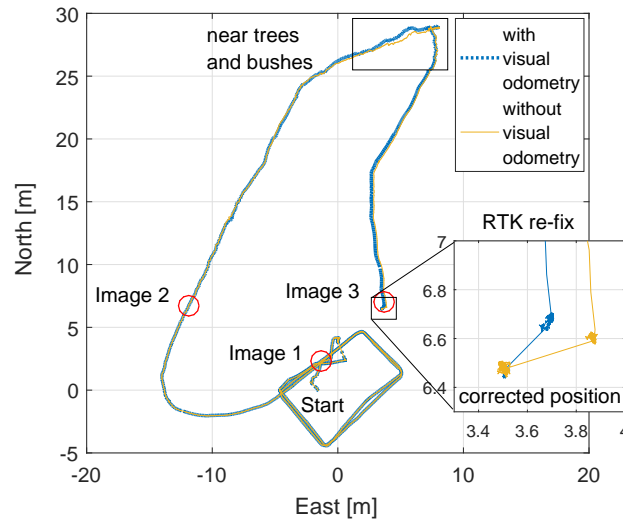


Fig. 8: Analysis of Multi-GNSS/ IMU/ wheel odometry tightly coupled RTK *with* and *without* integrated visual odometry.

Fig. 9 includes camera images with ~ 20 patch features on the grass at Pinakothek, Munich, with trees in the background. The illumination is higher in the left image than in the right image. The multilevel patch features are determined by ROVIO, and represented by squares. Green color denotes successfully tracked patch features and red color denotes rejected patches. The final (i.e. after iterative convergence) location of each landmark is shown with a small red dot surrounded by 4 green or red dots. The surrounding locations are checked for higher innovation residuals to keep (green) or reject (red) the patch features. The estimated uncertainty of each landmark location is shown by yellow ellipses. The largest uncertainty has the patch feature in the upper right part of the left image, where the image is very dark. We can observe that almost all patch features are in green, which indicates that grass patches can be tracked well.

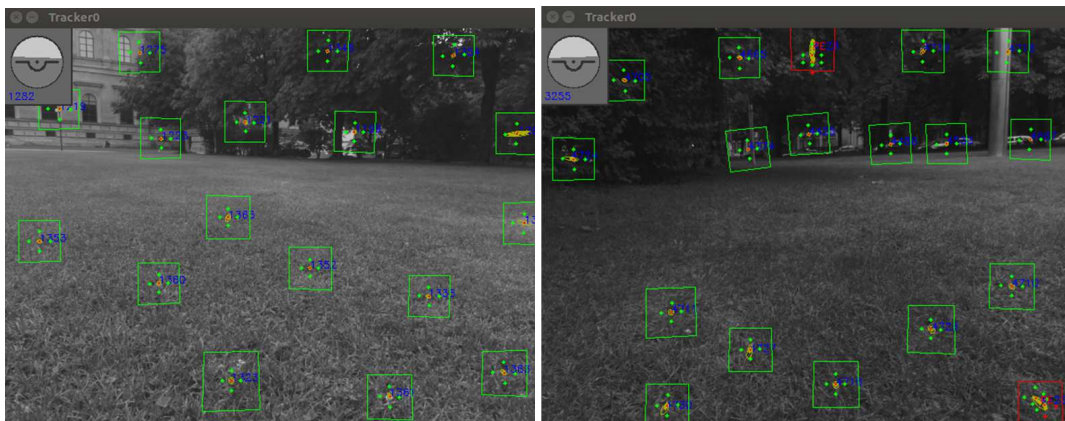


Fig. 9: Camera images with ~ 20 patch features on the grass at Pinakothek, Munich, with trees in the background.

