

# GPS and IMU Require Visual Odometry for Elevation Accuracy

Dirk Baumbach<sup>1</sup>, Hongmou Zhang<sup>1</sup>, Sergey Zuev<sup>1</sup>, Jürgen Wohlfeil<sup>1</sup>,  
Martin Knoche<sup>2</sup>, and Reinhard Klette<sup>3</sup>

<sup>1</sup> Institute of Optical Sensor Systems

German Aerospace Centre (DLR), Berlin, Germany

<sup>2</sup> N3T, Northland Innovation Centre, Whangarei, New Zealand

<sup>3</sup> School of Engineering, Computer, and Mathematical Sciences  
Auckland University of Technology (AUT), Auckland, New Zealand

## Abstract

Currently, self-localization of vehicles is primarily carried out by GPS receivers of different price and performance classes, which are permanently installed in vehicles. Known limitations of GPS-based localization include signal disturbances due to multipath propagation, as it may occur in narrow street canyons or tree alleys, as well as signal losses in tunnels or entrances to buildings, for example. It may happen that horizontal GPS channels still support a useful 2D positioning with low uncertainties while the GPS altitude channel has very large deviations or error uncertainties. Additional sensors (IMU or cameras) help to detect sections of trajectories where the GPS signal becomes unusable. Based on a traditional navigation strategy (i.e. integration of IMU and GPS), we also study the possible integration of visual odometry obtained from a stereo camera system. The paper reports about a large-scale project studying multi-sensor integration for very accurate and robust self-localization of vehicles.

## 1. Introduction

Logistical planning (availability, route planning, breakdown assistance) requires that companies know about the location of their vehicle fleets. For operating vehicles, it is often also appropriate to have even more accurate and robust self-localization data available for safety and efficiency reasons. Needs arise, for example, for automated driving in urban traffic with multi-level motorway junctions or in parking garages, but also, for example, for off-road truck driving, being one of the motivations for this paper. It is often insufficient to know just the 2D position (i.e. in the top-down view) of a vehicle relatively to a map; accurate elevation data are also needed.

The top-down view is presented in a Cartesian *East-*

*North coordinate system*, being a local planar approximation of geographic (spherical) coordinates. *Elevation* is measured with respect to the mean sea level, *height* with respect to a defined *starting point*; in our case we use the initial position of the test vehicle (see Fig. 1) when going on a test drive (to be more precise, the starting point is the position of the considered sensor at test-vehicle start).

Top-down views are the common way for representing trajectories (i.e. ignoring changes in the third dimension, the elevation dimension), even on a popular benchmark site such as [1]. A reason is that test sequences are typically provided for basically flat environments. In this paper we focus in particular on the third dimension, the height or elevation, thus also considering side-views of trajectories (i.e. East-height, or North-height diagrams).

Besides also taking elevation into account, we are also interested in comparing the performance of different sensors for self-localization of a vehicle. Currently, this task is primarily carried out by GPS receivers of different price and performance classes, which are permanently installed in the vehicles.



Figure 1: Used test vehicle on Otaika Valley Road. IPS setup during the measurement campaign. A GNSS antenna of a reference receiver is located left of the IPS-sensorhead.

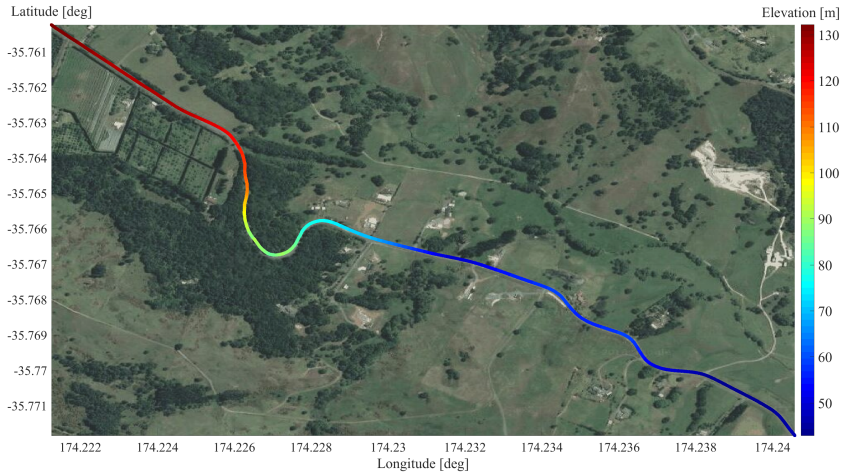


Figure 2: GPS-based trajectory (top-down view) for a 2.2 km-long section (7,608 recorded video frames) of the OVR. The shown trajectory (top-left to bottom-right) uses the color key shown on the right for encoding elevation along the trajectory.

For GPS it is known that the horizontal GPS channels may provide useful 2D positioning (i.e. in East-North coordinates) with low uncertainties while the GPS altitude channel has very large deviations or error uncertainties, thus producing unstable elevation estimates. Known limitations of GPS-based localization include signal disturbances due to multipath propagation, as they may occur in narrow street canyons or tree alleys, as well as signal losses in tunnels or entrances to buildings. We are also considering these GPS-limitations in our study, and ways for overcoming them.

The reported study is part of the research project “Digital Roads New Zealand”. The *Otaika Valley Road (OVR)* is of special interest (see Fig. 2); this road is  $\approx 12$  km long and considered to be dangerous due to frequent accidents, often involving timber trucks. Very accurate positioning tools may help to reduce the number of accidents.

There are three partners currently in the project “Digital Roads New Zealand”. The *Institute for Optical Sensor Systems* [2] at the *German Aerospace Centre (DLR)* developed an *integrated positioning system (IPS)* that may also be installed on a car for providing precise *odometry* (i.e. change of pose over time) of the vehicle on the road [3] while recording data by stereo vision and an *inertial measurement unit (IMU)*. Figure 1 shows an IPS installed on our test vehicle. The *Centre for Robotics & Vision (CeRV)* at *Auckland University of Technology (AUT)* contributes to computer vision solutions and dynamic visualisation technologies [4]. The *Northland Transport Technology Testbed (N3T)* aims at providing a test field for control and safety modules of vehicles [5].

Extensive data has been recorded and analysed within this project “Digital Roads New Zealand”. Thus, as a particular feature of this paper, we also state that our results are based on long-time recordings, by far exceeding test se-

quences currently used on [1]. However, in this brief conference paper we only address one test run (with a recording of 7,608 video frames) on OVR, illustrated in Fig. 2.

Figure 2 shows a 2D trajectory plot with a color-coded elevation profile using a low-cost GPS receiver. This aerial view of the OVR is given with geographic coordinates using a Google Maps excerpt. The elevation values are estimated altitude values (i.e. in relation to the mean sea level). From the starting point of the run at the top-left corner of the picture (elevation of about 130 m), the run proceeds over a distance of about 2.2 km with a height difference of over 80 m along the run; 40 m height difference appear in the shown small forest (of tall trees) within a very short distance of only about 800 m.

Our study shows that additional sensors (IMU or stereo cameras) are potentially able to improve elevation estimates when using an integrated approach; road sections where the GPS signal becomes unusable can be detected due to deviations from trajectories estimated by also using other sensors. Based on the traditional navigation strategy, defined by an integration of IMU and GPS, we also add the approach of *visual odometry (VO)* based on one stereo-camera system. Our study documents that a combination of *global* (i.e. IMU and GPS) and *relative navigation* (i.e. IMU and VO, as implemented in the IPS [6]) into one multisensor system can monitor situations where one of the two dominant sensors (GPS or VO) alone may malfunction or even fail.

The paper is structured as follows. Section 2 introduces the IPS. Section 3 discusses issues for GPS-only solutions. Section 4 reports about solutions using sensor integration. Section 5 concludes.

## 2. Integrated Positioning System

The used IPS (see Figure 1) is a low-cost stereo-vision-aided inertial navigation system which is developed by DLR. The IPS can measure the motion trajectory at 10 Hz in unknown indoor and outdoor scenes without any previous knowledge of the environment. The measured trajectory is a prerequisite for computer-vision projects such as 3D reconstruction, map-building, and so forth.

The IPS hardware includes a stereo-camera system (Prosilica GC1380H) and an IMU (ADIS-16488). The cameras may record CCD-progressive 1, 360 × 1, 024 stereo images at 30 fps (here used at 10 fps), with a focal length of 8.2 mm. For the IMU, both gyroscope and accelerometer have a bandwidth of 330 Hz, with a bias stability of 6.25°/h and 0.1 mg, respectively.

The IPS applies feature tracking for VO. A Kalman filter is used for filtering-out invalid features on dynamic objects. Subsequent pose updates, based on tracked valid features, are also guided by taking IMU measurements into account. For more details about the IPS, see [7, 8, 9, 10].

## 3. Global Positioning System

We discuss plots provided by *differential GPS* (dGPS) trajectories for the run shown in Fig. 2. Regarding the processing of GPS data, see [11]. We use a carrier-phase kinematic solution with processing of data at L1 and L2 frequencies. The used GNSS rover (i.e. receiver) was a Novatel OEM6 with active L1/L2 antenna, collecting both GPS and GLONASS satellite raw measurements for post-processing, using the reference GNSS station WNGR at Whangarei, at a distance of less than 10 km to the rover.

Figure 3 shows the variation in seen number of satellites during the considered run. For most of the run, the total number of seen satellites is more than ten. Significant are gaps when having tall trees on one or both sides of the road causing a loss of see satellites. The acquisition of two types of GNSS satellite data (here GPS and GLONASS) increase the total number and therefore the robustness of the solu-

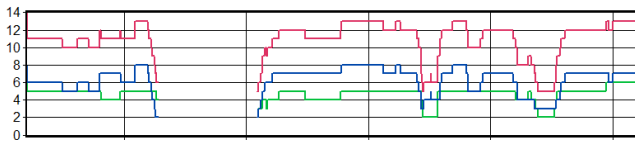


Figure 3: Number of satellites (ordinate values) seen by rover and by reference base station; abscissa shows GPS time during the drive (actual values are not of importance for this figure; difference between two vertical lines is 50 seconds). Color code: Total number of satellites is red, and in particular those of GPS in green, and of GLONASS in blue.

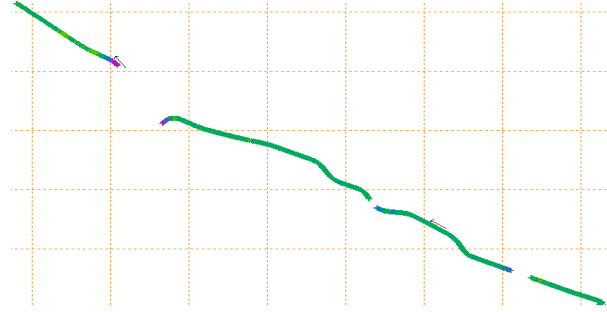


Figure 4: dGPS combined trajectory map (shown grid edges are 300 meters) of forward and reverse dGPS solution. Trajectory color turns from green to dark blue and even pink if number of satellites drops.

tion. Using just GPS satellites would be near to the critical minimal number of four satellites, what would cause a poor and unstable dGPS solution.

Obviously, there is a significant data gap for available satellite data due to satellite shadows caused by trees near to the road. At those parts, the dGPS may fail at initialisation to resolve ambiguities, and has to start an automatic real-time kinematic initialisation during the drive. This reduces position accuracy around those gaps; see Fig. 4.

We processed a dGPS solution in offline mode for producing a reference trajectory of the rover with high accuracy [11]. Differential GNSS is a technique which allows us an improvement of the performance of GNSS positioning. The main principle is to calculate the position of the receiver on the rover relatively to another receiver, called the *base station*, for which we know precise coordinates.

This allows us to cancel out main GNSS error components caused by ionosphere delay or clock drift, since they are common for rover and base station. The distance rover-to-base is therefore a limiting factor which has been taken into account for the planing of the measurements.

Carrier-phase measurements are a very precise technique for differential measurements; they allow us to estimate positioning at cm-level accuracy. This technique measures the distance from receiver to the satellite in units of circles of carrier frequency, and requires a special carrier phase receiver. The processing can be done offline.

For the discussed measurement campaign, our rover was equipped with a second high quality GPS receiver Novatel OEM6 with active L1/L2 antenna (see Fig. 1). During the drive, simultaneously to the IPS system, the reference receiver collected GPS and GLONASS satellite raw measurements with L1 and L2 frequency for post-processing.

Results of the DGNSS post-processing for the considered OVR run are represented in Figs. 5 to 7. Significant are some data gaps in the DGNSS solution related to segments of the run with tall trees near to the road. See especially

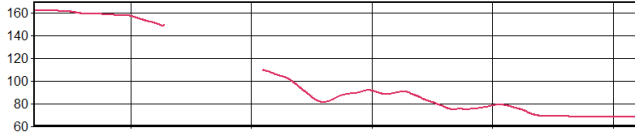


Figure 5: Altitude (ordinate values in meters) profile plot in pink over Ellipsoid-WGS84; abscissa shows GPS time during the drive

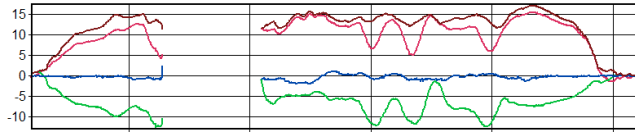


Figure 6: Velocity profile plot (ordinate values in meters per second; abscissa shows GPS time during the drive). Color code: East in pink, North in green, up in blue, and horizontal speed in dark red

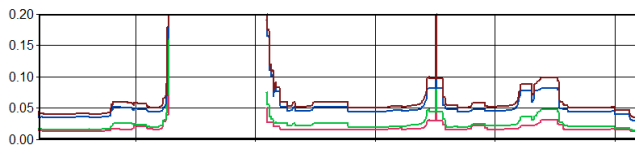


Figure 7: Estimated position accuracy plot. For most of the trajectory, the dGPS solution provides high absolute positioning accuracy, with an error below 5 cm. The ordinate values are the standard deviation in meters; abscissa shows GPS time during the drive. Color code: East pink, North light green, height in dark blue, and trace (i.e. mean standard deviation in 3D) in dark red

Fig. 5 for missing altitude estimates. Regarding Fig. 6, start and end position of trajectory was operated as static (i.e. zero velocity for 10-15 seconds) for a stable initialization of the dGPS solution.

After the shown gaps (at tall trees), a re-initialization of dGPS solution must be applied ‘on-the-drive’ which causes an estimated error of up to 10 cm in horizontal direction, and up to 30 cm for height; see Fig. 7.

Figure 8 shows a longitude-elevation diagram and the associated standard deviation of the GPS receiver, on the same longitude coordinate axis as in Fig. 2.

This shows again the unfavorable GPS reception conditions within the forest section of the road. While at the beginning of the ride the standard deviation of the vertical channel is  $\approx 1.2$  m only, it increases strongly when driving through the forest section (to  $\approx 6.5$  m), and remains very large until leaving the forest ( $> 2.5$  m). At the same time, the elevation decreases sharply. The graph shows an up and down of  $\approx 5$  m in a distance of 10 m only, which is, obviously, implausible for a road.

The upper photo in Fig. 8, recorded by the camera system, indicates a rather monotonous descent of the road. In

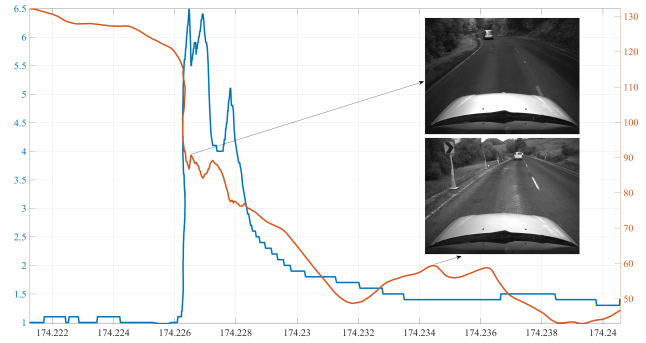


Figure 8: Longitude-elevation diagram (red curve, in meters; see ordinate on the right) along the GPS-based trajectory shown in Fig. 2; note the identical scaling of the longitude coordinates (the abscissas) in degrees. The diagram also shows the standard deviation of measured elevation (blue curve, in meters; see ordinate on the left), and two examples of recorded frames

summary, such a narrow and winding road, with many tall trees left and right of the road, has a very negative impact on GPS reception. In the following section of the road, with only a few trees and free views to the satellites (see lower image in the figure), the standard deviation returns to the usual level of  $\approx 1.5$  m. The focus of our research is on the usability of very unstable GPS measurements combined with other sensors. In particular, the altitude channel of the GPS receiver appears to be problematic in the forest section of the road.

## 4. Integrated Solutions

Figure 9 shows the normalized standard deviation of 3D positions for GPS (blue line) and different integrative approaches. Note that IMU+CAM+GPS is superior to others.

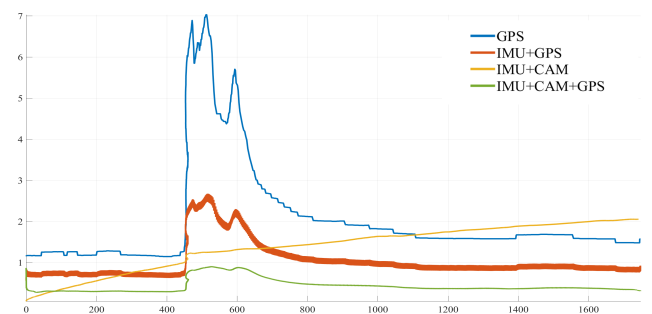


Figure 9: Normalized 3D position standard deviations (ordinate values in meters; abscissa for East values in meters) for different sensor options. Color code: GPS in blue, IMU and GPS in red, IMU and stereo camera in orange, and all three (IMU, GPS and stereo camera) in light-green

The standard deviation constantly increases for IMU+CAM because of a missing absolute guide. The use of GPS alone leads to a discontinuity at the critical place, but stabilizes soon again. The normalization of the standard deviations was performed for purposes of clarity and because the amplitude of position uncertainties is primarily determined by the altitude channel.

Similarly to Fig. 8, a strong increase in standard deviation for GPS-only position estimation appears when driving through the forest.

The uncertainty for the IMU plus camera combination (orange line) starts near zero because the local navigation knows no uncertainties at the beginning. The 3D position uncertainty is unlimited in the further course, that is, it increases depending on scenes over time, based on the exclusively relative support from visual odometry.

By extending GPS-only signal analysis with an IMU (red line) and a camera component (green line), the system uncertainty can be lowered repeatedly. This is a general observation in our experiments: An extension of an GPS plus IMU solution by also using relative visual odometry measurements (in our experiments based on a stereo-camera system) can help to detect situations of poor or even missing GPS measurements, and to improve the plausibility and robustness of the obtained navigation solution.

Figures 10 and 11 show trajectory results for the discussed different combinations of sensors when projected either into the common top-down (i.e. North-East) view or into one of the orthogonal side-views.

In Fig. 10, curves for IMU and GPS, or for GPS alone are rarely visible because they are covered by the light-green curve (when using all three sensors); the dotted orange curve (IMU and stereo cameras) deviates a little from the light-green curve. This demonstrates a dominant consistency between all the estimated 2D trajectories.

Regardless of the GPS raw data signal (blue dots), additional horizontal or vertical GMES points (shown in black)

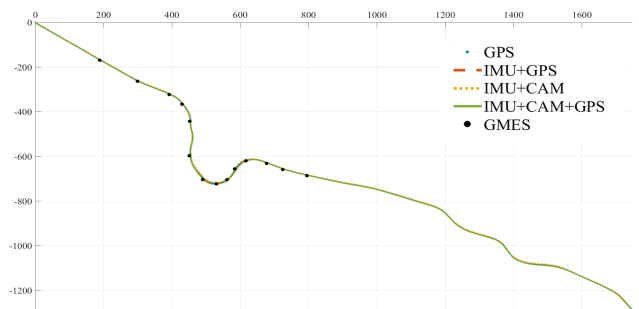


Figure 10: North-East (i.e. top-down view) trajectories, using the same color code as in Fig. 9, and discrete GMES locations (black dots); abscissa for East and ordinate for North, both values in meters. Driving direction is basically left to right

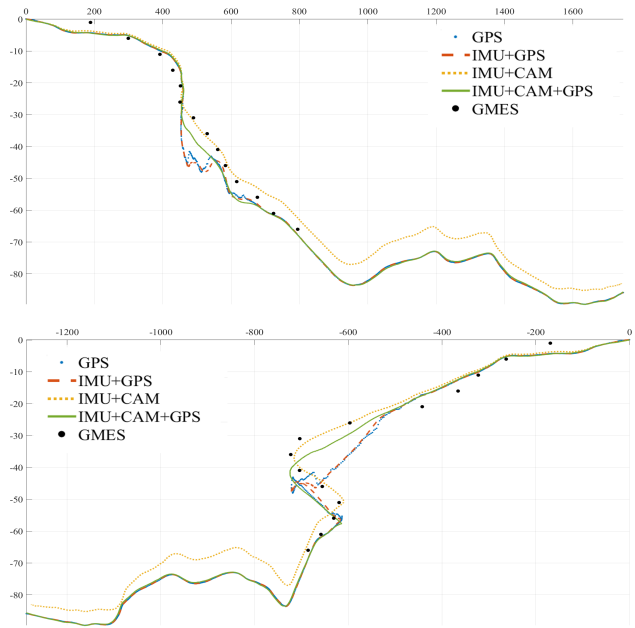


Figure 11: *Top*: Height-East (i.e. one of two orthogonal side views) trajectories using the same color code as in Fig. 10; abscissa for East and ordinate for Height, both values in meters. Driving direction is basically left to right. *Bottom*: Height-North (i.e. the second orthogonal side view) trajectories using the same color code as in Fig. 10; abscissa for North and ordinate for Height, both values in meters. Driving direction is basically right to left.

were obtained from the *Google Map Elevation Service* API to serve as a quantitative reference.

While the top view in Fig. 10 shows a nearly identical course of GPS raw data measurements (blue dots), different sensor fusion strategies (red, orange, or green line), and the GMES reference points (black), differences in both height maps in Fig. 11 are clearly visible.

The GPS signal (see blue dots at East = 500 m and Height = -45 m in Fig. 11 top view) shows large height differences within the forest region (refer to Fig. 8), relatively to the distance traveled. The fusion of GPS with the IMU (red dashed line) leads to a kind of smoothing of the trajectory; the basic course cannot be changed.

The local navigation solution, consisting of IMU and VO only (orange dotted line), shows an increasing drift in all three views, compared to the GPS raw data measurement, where the drift in the vertical channel is most significant. We emphasize the continuous descent within the forest section (Fig. 11); this appears to be plausible for the given road segment. The term *local* means here that the IMU+CAM solution was calculated in a coordinate system corresponding to the IMU system at time zero. In order to compare this calculated trajectory with those calculated when using a global sensor, it was manually rotated around the *Z*-axis

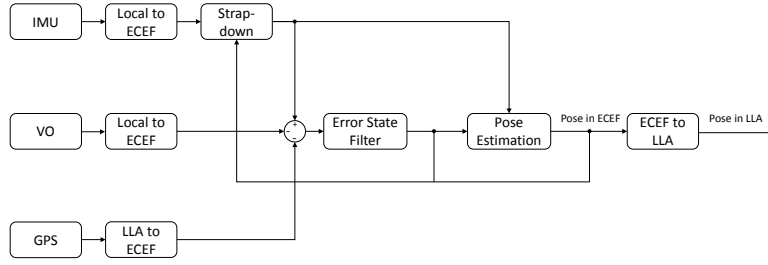


Figure 12: Sketch for IMU, VO, and GPS integration

such that it corresponds with the North-East convention.

Figure 12 shows a sketch of performed data fusion. Because IMU, VO, and GPS data are recorded in different coordinate systems, first all the data are transformed into *Earth-centered, Earth-fixed* (ECEF) coordinates. Original GPS measurement is in *latitude, longitude, and altitude* (LLA) coordinates, to be transformed by

$$\begin{bmatrix} \bar{p}_x \\ \bar{p}_y \\ \bar{p}_z \end{bmatrix} = \begin{bmatrix} r_s \cos \lambda_s \cos \iota + h \cos \mu \cos \iota \\ r_s \cos \lambda_s \sin \iota + h \cos \mu \sin \iota \\ r_s \sin \lambda_s + h \sin \mu \end{bmatrix} \quad (1)$$

where  $\mu$  is latitude,  $\iota$  is longitude, and  $h$  is altitude;  $\lambda_s$  and  $r_s$  are defined as follows:

$$\begin{aligned} \lambda_s &= \arctan((1-f)^2 \tan \mu) \\ r_s &= \sqrt{\frac{R^2}{1 + (1/(1-f)^2 - 1) \sin^2 \lambda_s}} \end{aligned} \quad (2)$$

where  $f$  defines flattening and  $R$  is the equatorial radius of the Earth.

By taking differences between VO and IMU, or GPS and IMU as input, the error state of the system can be updated by an error state Kalman filter. The optimized system pose is finally obtained by adding the filtered error state to the strapdown output. At the end, the coordinate of the system pose is transformed back into LLA coordinates to compare with other benchmarks.

The height images in Fig. 11 clearly show the different course of a combined trajectory compared to that provided by the GPS signal only; it is also steadily descending and has no jumps. As an attempt of quantitative support for the accuracy of the IMU+CAM+GPS solution, we plotted GMES elevation values as available in the area of the forest at a distance of 5 m at most. We also notice a steadily declining course. This points again to erroneous GPS altitude signals in this area.

## 5. Conclusions

The paper addressed the problem of elevation (or altitude or height) estimation in self-localization, typically overlooked by discussing 2D trajectories only. Extensive experiments (here only illustrated for one segment of the Otaika

Valley Road) indicate that under good viewing conditions, an integrative use of visual odometry can solve issues otherwise given if GPS and IMU are the only used sensors. The paper also documents a need to include elevation-dynamic test sequences into common odometry benchmarks.

## References

- [1] Visual Odometry. The KITTI Vision Benchmark Suite, [http://www.cvlibs.net/datasets/kitti/eval\\_odometry.php](http://www.cvlibs.net/datasets/kitti/eval_odometry.php), June 2018.
- [2] DLR Institute for Optical Sensor Systems. [www.dlr.de/os/en/](http://www.dlr.de/os/en/), 2018.
- [3] A. Börner, D. Baumbach, M. Buder, A. Choinowski, I. Ernst, E. Funk, D. Grießbach, A. Schischmanow, J. Wohlfeil, and S. Zuev, “IPS – A vision-aided navigation system,” *Advanced Optical Technologies*, 6(2):121–129, 2017.
- [4] Centre for Robotics & Vision. [cerv.aut.ac.nz](http://cerv.aut.ac.nz), 2018.
- [5] The Northland Transport Technology Testbed. [www.n3t.kiwi](http://www.n3t.kiwi), 2018.
- [6] D. Grießbach, “Stereo-vision-aided inertial navigation,” PhD thesis, Freie Universität, Berlin, 2014.
- [7] D. Grießbach, D. Baumbach, and S. Zuev, “Stereo-vision-aided inertial navigation for unknown indoor and outdoor environments,” *Proc. Int. Conf. Indoor Positioning Indoor Navigation*, IEEE Xplore, 2014.
- [8] H. Zhang, D. Grießbach, J. Wohlfeil, and A. Börner, “Uncertainty model for template feature matching,” *Proc. Pacific-Rim Symposium Image Video Technology*, LNCS 10749, pp. 406–420, 2017.
- [9] H. Zhang, J. Wohlfeil, and D. Grießbach, “Extension and evaluation of the AGAST feature detector,” *Proc. ISPRS Congress Annals*, vol. 3, pp. 133–137, 2016.
- [10] H. Zhang, J. Wohlfeil, D. Grießbach, and A. Börner, “Eligible features segregation for real-time visual odometry,” *Proc. 3D-NordOst*, in print, 2017.
- [11] M. S. Grewal, L. R. Weill, and A. P. Andrews, “Global Positioning Systems, Inertial Navigation and Integration,” John Wiley & Sons, pp. 190–235, 2007.