Precise Attitude and Position Determination of the Trailer using a Single Camera System for Agricultural Applications

Himanshu Sharma, Arkadiusz Szumski, Mathias Philips-Blum,
Prof. Thomas Pany, Prof. Bernd Eissfeller
Institute of Space Technology and Space Applications
Universität der Bundeswehr, Munich, Germany

BIOGRAPHY (IES)

Himanshu Sharma is a research associate at the Institute of Space Technology and Space Applications (ISTA). He received his Bachelors in Technology (B. Tech) degree in the field of Electronics and Communication from maharishi Dayanand University, Rohtak, India. He accomplished his Masters in Science (M.Sc.) in the field of Comminutions and Signal Processing from Technical University of Ilmenau, Thuringia, Germany. His research interests are Digital Signal Processing, Image processing. He has been involved in the projects that focus particularly on Attitude determination using the inertial sensor system.

Arkadiusz Szumski is a research associate at the Institute of Space Technology and Space Applications (ISTA). He received his engineer diploma in space telecommunication systems from Telecom Bretagne, France and a M.Sc. degree in radio-communication from Gdansk University of technology, Poland. His research interests are signal processing algorithms. He has been involved in projects that focus particularly on frequency and phase synchronization, interference mitigation and lately also on inertial sensors based navigation systems.

Mathias Philips-Blum received his master in Geodesy and geoinformatics from the Leibniz University of Hannover, Germany. Since 2017 he is a research associate at the University of Federal Armed Forces Munich at the Institute of Space Technology and Space Applications. He currently works in the field of sensor fusion and navigation.

Prof. Thomas Pany is with the Universität der Bundeswehr München at the faculty of aerospace engineering where he teaches satellite navigation. His research includes all aspects of navigation ranging from deep space navigation over new algorithms and assembly code optimization. Currently he focuses on GNSS signal processing for Galileo second generation, GNSS receiver design and GNSS/INS/LiDAR/camera fusion. To support this activities, he is developing a modular GNSS test bed for advanced navigation research. Previously he worked for IFEN GmbH and IGASPIN GmbH and is the architect of the ipexSR and SX3 software receiver. He has around 200 publications including patents and one monography.

Prof. Bernd Eissfeller is a full professor of navigation and director of the Institute of Space Technology and Space Applications at the Universität der Bundeswehr München. He is responsible for teaching and research in navigation and signal processing. Until the end of 1993 he worked in industry as a project manager on the development of GPS/INS navigation systems. He received the Habilitation in Navigation and Physical Geodesy, and from 1994-2000 he was head of the GNSS Laboratory of the Institute of Geodesy and Navigation. From 2000 to 2008 he was Vice Director of the IGN until he became the Director of the Institute which was formerly led by Prof. Günter Hein. Eissfeller is the author of more than 300 scientific and technical papers.

ABSTRACT

The emerging use of GNSS (satellite navigation) is helping farmers to achieve high productivity without making compromise with the quality. Further development of the farming system could improve the contemporary agriculture. It is still possible by improving the precise placement of the seeds, ploughing, and harvesting down to the centimeter level accuracy.

The present precise agriculture navigation systems are mostly limited to precise navigation of the main components, like tractors and combine harvesters. In the state of art technique, DGPS (Differential GPS) or RTK (real-time-kinematics) systems and inertial sensors are used to determine the precise relative position and attitude of the main driver. Precise seeding is more complex due to the complexity of the multicomponent machines. Assuring the positioning accuracy of any attached component to a tractor machine would require installation of an additional DGPS or RTK unit, which because of the high cost factor, is not the optimum solution.

The approach presented in this paper demonstrates the use of a camera system to measure the precise relative position and, additionally, attitude of the trailer with respect to the tractor. The researched navigation system is suited for a trailer pulled by a tractor. In our test setup, the role of a tractor is perceived by a Volkswagen T5 van and the trailer is demonstrated by commercial cargo trailer.

The paper also discusses the various approaches investigated for the development of the algorithm. The reader will have an insight into the challenges related to the attitude determination (van-trailer scenario) with in-depth analysis of the algorithm. The testing environment including testing procedure, results and comparison between attitude measurements using proposed algorithm and multiple DGPS/RTK and geodetic instruments will be discussed in details. The paper will conclude the future possibilities and the extension of the work to make it more precise and reliable to be ready for commercialization.

INTRODUCTION

The commercial positioning and navigation solutions available in the market are highly dependent on satellite availability. During bad weather, less satellite visibility decorate the positioning solution. Inertial navigation systems are one of the alternate solutions but the cost for such solution is sky high. On the other hand, farmers now a days try to cope up with the technology trends and are open to the solutions which could save their efforts and time. Seeding or sowing is a very important and critical step in farming.

To perform better seeding, we must ensure proper distance between seed because if plants are overcrowded, they will not to get enough water, nutrients and sunlight, resulting in yield loss. If they are planted too far from each other, valuable land is left unused. The mechanical technique used for loosening the soil is also very time consuming. Precise loosing can help farmer to save time and fuel over a larger period of time. In order to meet this vision, the farmers shall be equipped, among the others, with an appropriate navigation system for the mobile machines. At the same time an economical aspect must be considered. It means in a consequence that the applied navigation system should be easy to use, should make it possible to save time and ultimately should be a part of an autonomous agriculture complex.

Presented research targets for 10 cm accuracy of a seeding machine coordinates, assuming the distance from the host machine 3 meters. The absolute position of fix point is known from tractor RTK/INS system. The relative host to seeding machine attitude accuracy required for the particular application is about 2 degrees.

The basic idea behind the development of the algorithm is to determine the attitude and the position of the trailer using the images from a vision camera. This is possible thanks to the relative navigation w.r.t. the reference RTK/INS system mounted on the tractor. The reference RTK system combined with inertial sensors on the tractor does provide the position and attitude necessary. From the knowledge of the precise position and attitude of the tractor, the trailer is further observed from the tractor perspective through the camera system.

Figure 1 : Tractor and Trailer in the Field

The analysis depicted that the trailer can change the attitude freely in any axis with respect to fixing point on the tractor. The attitude of the trailer with respect to the tractor has been measured using a Giga Ethernet (GigE) camera. The development of the algorithm is performed using OpenCV platform. The algorithm includes the feature extraction using the camera and the frame transformation to achieve the attitude determination w.r.t to the reference points of the tractor and the trailer. The preliminary focus of feature extraction work is to read the images and localize the position of the trailer on the image [1]. The image in 2D is then processed into 3D attitude determination.

Since, the movement of trailer w.r.t tractor is free along fixing point (hook) on the tractor, we also need a mathematical model to converge from orientation in camera frame to orientation in hook frame to determine orientation of trailer w.r.t tractor. Where, **C** is direction cosine matrix, **t** is the reference point inside trailer, **h** is the hook point on tractor (joining of hook and trailer), **r** is the reference point on the trailer used for image processing.

$$C_h^t = C_r^t \times C_h^r \qquad [1]$$

Direction cosine matrices of reference point **r** on the trailer w.r.t hook is difficult to calculate due to free motion of the trailer. To determine the attitude of the trailer with respect to the reference point on the tractor, we must performed the frame transformation from the camera to the reference point on the tractor. Processing must be sensitive enough to assess the dynamic motion of the tractor and the trailer. The vibrant environment is one of the noise components, which is also investigated.

The practical testing of the algorithm has been performed constructing similar relative motion scenario. Test drives have been performed with complete setup in the premises of Universität der Bundeswehr Munich. The camera is fixed on the rear part of the van to have a clear view of trailer.

Figure 2 : Van and Trailer to replicate tractor and trailer scenario

To assess the accuracy of our low-cost trailer positioning system, we equip the trailer with a second RTK system and compare the vision based solution with the RTK solution. We focus on different test cases including static measurements and measurements for typical farming trajectories. The accuracy is expected to depend on the movement pattern due to unavoidable latencies and limitations of the vision based solution.

LINEAR N-POINT CAMERA POSE ESTIAMTION

Definition of Pose Estimation

Space resection or **Pose estimation** is the process of determining the position and orientation of the calibrated camera with respect to the known reference points [3]. The pose can be changed either by moving the camera w.r.t object, or moving the object w.r.t camera. The goal of the pose estimation also referred as **Perspective-n-Point** problem is to find the pose of an object when we have a calibrated camera, and we know the location of **n** 3D points on the object and the corresponding 2D projections in the image [2].

A 3D rigid object is restricted to two types of motion with respect to camera [4].

1. **Translation Motion:** Shifting or moving the camera from current 3D location (X, Y, Z) to another location (X', Y', Z'). The translation vector denoted by vector **t** is (X'-X, Y'-Y, Z'-Z).
2. **Rotation Motion:** The other motion termed as rotation motion represents the rotation of camera or the object. Rotation motion can be represented in several ways including **Euler angles** (roll, pitch and yaw), a **3 x 3** rotation matrix, or a direction of rotation and angles.

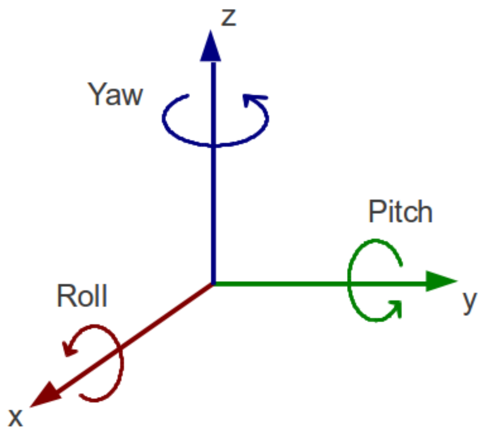

Figure 3 : Roll, Pitch and Yaw motion

Pose Estimation mechanism

In order to perform pose estimation of an object, we need the following inputs.

1. **2D coordinates of a few points:** 2D coordinates points is the 2D (X, Y) location of few points in the image.
2. **3D locations of the same point:** 3D location of similar 2D coordinates in some arbitrary reference frame (World Coordinates). Here the 3D coordinates position does not corresponds to some 3D model for the image.
3. **Intrinsic camera parameters**: Focal length of the camera, optical center in the image and the radial distortion parameters comes under the category of intrinsic camera parameters.

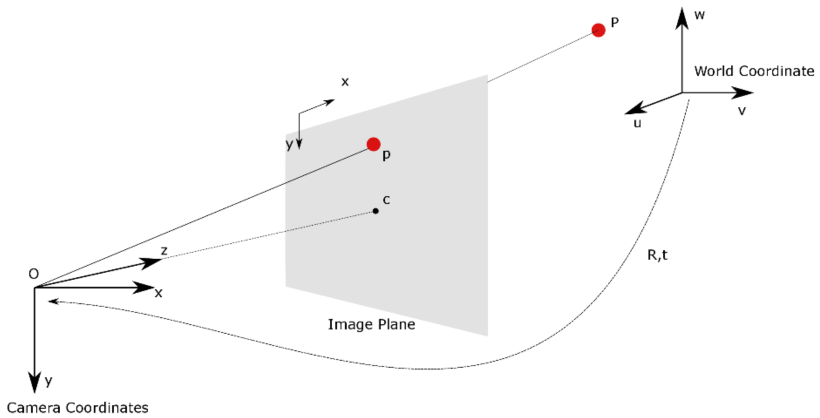

Figure 4: Coordinates Frame Description [5]

In order to better understand the mechanism for Pose estimation consider figure [4], with O as a center of camera. We are intended to compute the equations for projection p (image plane) of 3D point P in world coordinate. Let's us assume we know the coordinate of location (U, V, W) of a 3D point P in world coordinates. If we know the rotation matrix **R** and translation vector **t**, we can calculate the location (X, Y, and Z) of the point P in the camera coordinate system using the equations [2-4].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \boldsymbol{R} \begin{bmatrix} U \\ V \\ W \end{bmatrix} + \boldsymbol{t} \tag{2}$$

$$= \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [\boldsymbol{R} \mid \boldsymbol{t}] \begin{bmatrix} U \\ V \\ W \end{bmatrix} \tag{3}$$

$$= \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} r_{00} & r_{01} & r_{02} & t_x \\ r_{10} & r_{11} & r_{12} & t_y \\ r_{20} & r_{21} & r_{22} & t_z \end{bmatrix} \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix} \tag{4}$$

In order to measure the rotation and translation vector we must know (X, Y, Z). Since we have no information regarding the (X, Y, Z) in camera coordinate frame, we must use direct linear transform by using available 2D coordinates and intrinsic camera parameters [5].

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = s \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{5}$$

Where, f_x and f_y are the focal lengths of camera in x and y directions, andc_x, c_y is the optical center. s is the scaling factor used since, the depth of image is unknown. With all the available inputs the above equation can be solved for (X, Y, Z) and finally rotation and translation vectors.

IMPLEMENTATION

Hardware and Software Setup

The camera used for the purpose is a Giga Ethernet VLG-20C.I, with the resolution of 1624 x 1228 pixels. The camera operates at the maximum of 27 fps with the operation voltage ranges from 12-24V DC. The detailed technical description about the camera are in the table [1].

Tools	Platform
Development Environment	Visual C++ 2012
OS	Windows 10, x64
Image Processing	OpenCV 2.4.13
SDK	Baumer GAPI2 (windows)
Gig-Ethernet Camera	• Baumer VLG-20C.I [5] • Resolution : 1624 X 1228 pixels • Maximum 27 fps • Operating voltage : 12-24V DC

Table 1: Description of hardware and Software Setup

Figure 5 : Baumer GigaE Vision Camera

The camera has been mounted on the rear top of the van to maximize the visibility of trailer. The Image acquisition and Image processing algorithm is developed on c++ platform using OpenCV.

Image Acquisition and Image processing

The image acquired from the image acquisition step is now processed to determine the orientation of the image in camera frame. In order to determine the extrinsic camera parameter, camera needs to be calibrated to minimize distortions. Radial and tangential distortions are the most common types of distortion in cameras.

Radial Distortion: This distortion as we move away the center of camera. The effect of such distortion is that the straight lines appear to be curved one.

Tangential Distortion: This time of distortion arises in case the image taking lens in not aligned parallel to the image plane. This results in some area of image appearing to be closer than actual.

Further, we need to obtain intrinsic camera parameters which are specific to a camera. This includes information related to focal length (fx, fy), optical center (cx, cy) etc. It is also termed as camera matrix as shown in equation [6].

$$camera\ matrix = \begin{bmatrix} fx & 0 & cx \\ 0 & fy & cy \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$

To determine the extrinsic camera parameters is also called pose estimation. In order to do pose estimation, we have to provide some sample image of a well-defined patter (chess board). We find some specific points in it e.g. corners. We know its coordinate in real world space and we know its coordinate in image. Considering chessboard was kept stationary (z=0) the 3D points (object points) can be written as (0, 0, 0), (1, 0, 0), (2, 0, 0) for (x, y, z).

In the next step we have to find chessboard pattern using function **findchessboardcorners ()**. This function detects all the corners of the chessboard and gives us 2D image point's matrix. Using 2D and 3D points function **solvePnP ()** provides rotation and translation vector of the image as the output.

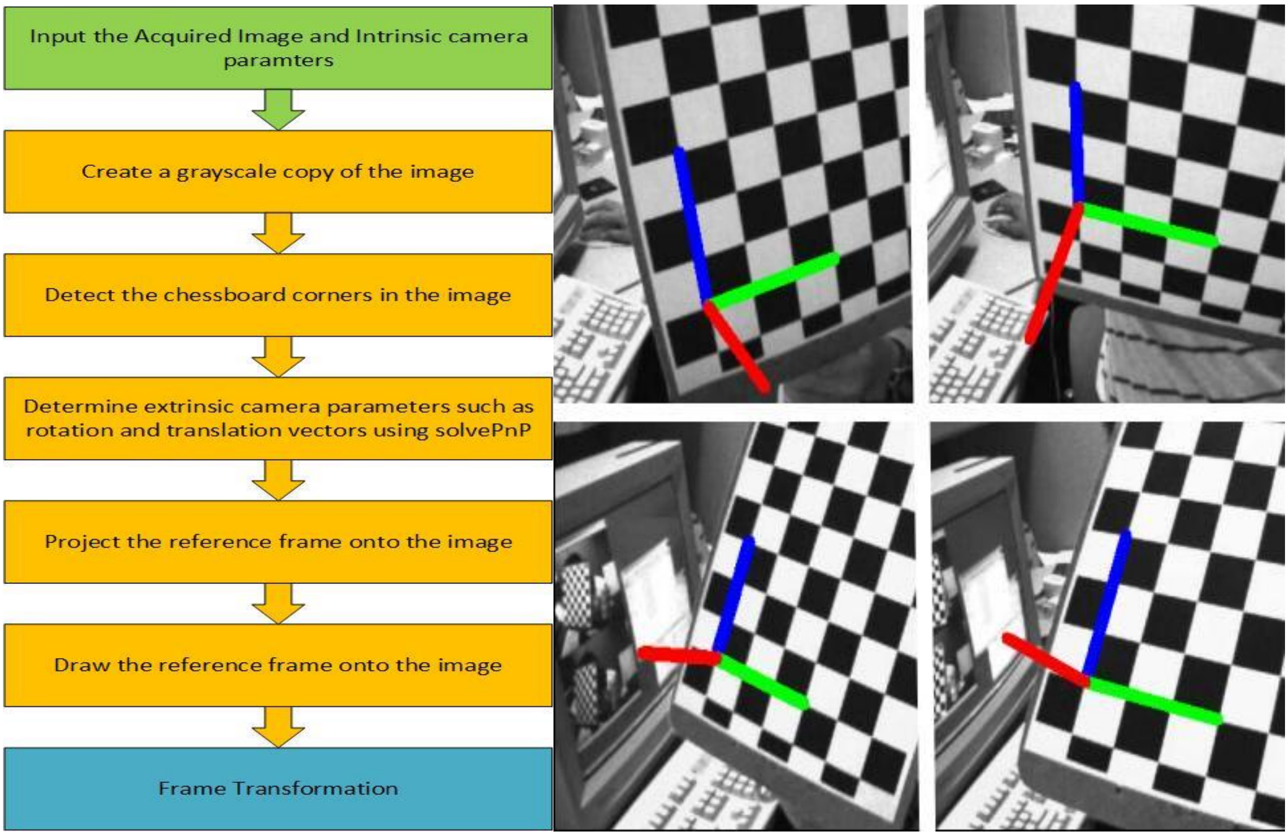

Figure 6: (Left) Design Flow for Image processing Algorithm, (Right) Reference frame projection on chessboard [4]

We want to draw our 3D coordinate axis (X, Y, Z axes) on our chessboard's first corner. X axis in blue color, Y axis in green color and Z axis in red color. So in-effect, Z axis should feel like it is perpendicular to our chessboard plane. Now let's create a function, draw which takes the corners in the chessboard (obtained using findChessboardCorners ()) and axis points to draw a 3D axis. [1]

Then as in previous case, we create termination criteria, object points (3D points of corners in chessboard) and axis points. Axis points are points in 3D space for drawing the axis. We draw axis of length 3 (units will be in terms of chess square size since we calibrated based on that size). So our X axis is drawn from (0, 0, 0) to (3, 0, 0), so for Y axis. For Z axis, it is drawn from (0, 0, 0) to (0, 0, -3). Negative denotes it is drawn towards the camera.

Frame Transformation

The goal of the algorithm is to determine orientation and position of the trailer w.r.t the bus. It is represented with the direction cosine matrix (DCM) between the t-frame and h-frame. This calculation is not possible to be performed in a direct relation. Therefore two additional frames, reference(r) and camera (c) have been introduced as shown in figure [7].

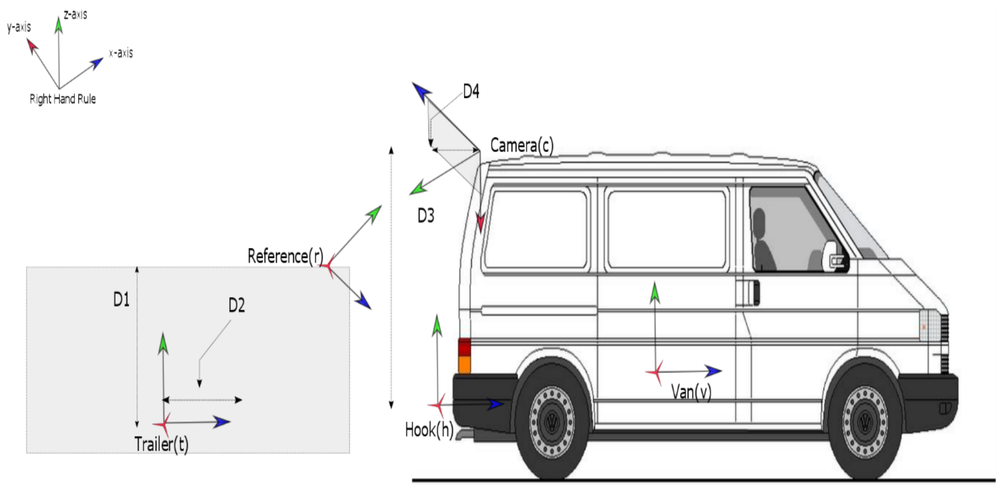

Figure 7: Frame representation

<u>Illustration of frames</u>

Frame t: frame with point trailer as origin
Frame r: reference frame of chessboard
Frame c: camera frame
Frame h: hook frame
Frame v: van frame

Origin of the trailer frame (t) is located at the center of the sensor plate installed inside the trailer. Expressing mathematically, the direction cosine matrix for the orientation of frame t w.r.t the frame h can been written as [7].

$$C_h^t = C_r^t \times C_h^r \qquad [7]$$

Due to the geometrical limitation, it is not possible to determine the Orientation of frame r w.r.t the h frame. The idea developed here, is to compute Orientation of frame r w.r.t to frame h using camera frame C. The equation can now be represented as [8].

$$C_h^t = C_r^t \times C_c^r \times C_h^c \qquad [8]$$

Direction cosine matrix for frame t w.r.t frame r is based on the orientation determined using Euler angle rotation roll (), pitch (), and yaw () as represented in equation [9].

$$C_r^t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\varphi & -sin\varphi \\ 0 & sin\varphi & cos\varphi \end{pmatrix} \times \begin{pmatrix} cos\theta & 0 & sin\theta \\ 0 & 1 & 0 \\ -sin\theta & 0 & cos\theta \end{pmatrix} \times \begin{pmatrix} cos\emptyset & sin\emptyset & 0 \\ sin\emptyset & cos\emptyset & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{for the X-Y-Z transformation} \qquad [9]$$

Direction cosine matrix for the reference frame w.r.t camera frame is dynamic and it changes with every processed image. All other DCMs are constant and express the geometry of the fixing camera or the reference points on the trailer.

RESULTS

Measurement -Van Setup for Reference Attitude Measurement

In order to compare the attitude accuracy of the camera system, a reference system was used. The reference system works with the Theodolite as shown in figure [8].

Figure 8: (Left) Hardware setup for the measurement (right) theodolite setup to obtain reference orientation

Hardware	Location
GNSS Receiver	Roof of the trailer, van
Camera	Roof of the van
Chessboard (9 X 6 crossings)	Front of the trailer
Inertial Meas. Unit	Inside Van

Table 2: Description of hardware placement

The coordinate points with hook as reference point were then converted to obtain model of trailer w.r.t van and finally calculate the orientation of trailer w.r.t van.

In order to testify the solution, trailer w.r.t van was setup in different positions and static test were performed. The detailed description about the test scenarios can be found in the table [3].

Scenario	Position
Scenario #1	Trailer – Van inline
Scenario #2	Trailer Position left to van
Scenario #3	Trailer Position left to van
Scenario #4	Trailer left side inclined
Scenario #5	Trailer right side inclined
Scenario #6	Trailer lifted

Table 3: Description of the various testing scenarios

Testing measurements analysis

In order to analyses the influence of external disturbances on the measurements and precision of the attitude determination, 270 consecutive measurements were compared within different trailer-van settings each of the image processed delivered an attitude of the trailer. The results are presented on the histograms in a form of transformation angles, roll, pitch and yaw. The histograms for each scenario and for each transformation angle was compared to the normal distribution making it easier to rank the performance of the attitude determination. Distributions is centered to mean values depicted by red line (--). All measurements have very similar output. The histograms of two of the scenarios are presented in the figures [9-10] below.

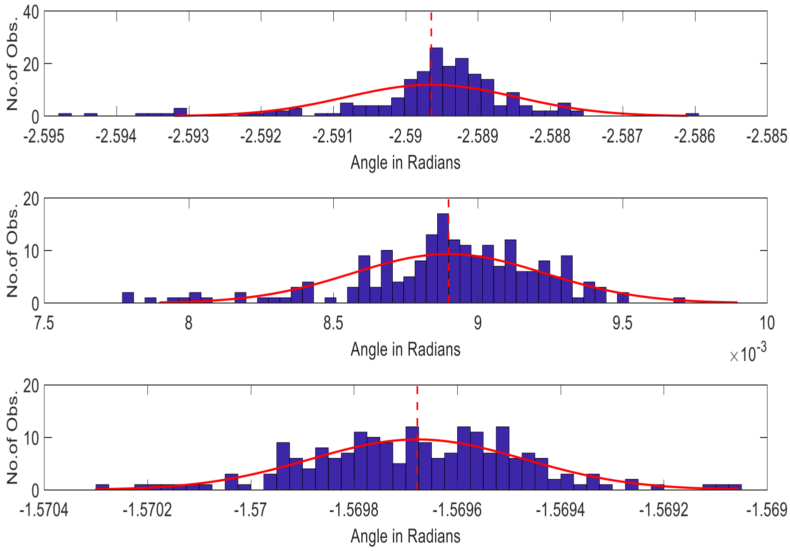

Figure 9: Trailer and Van in -line

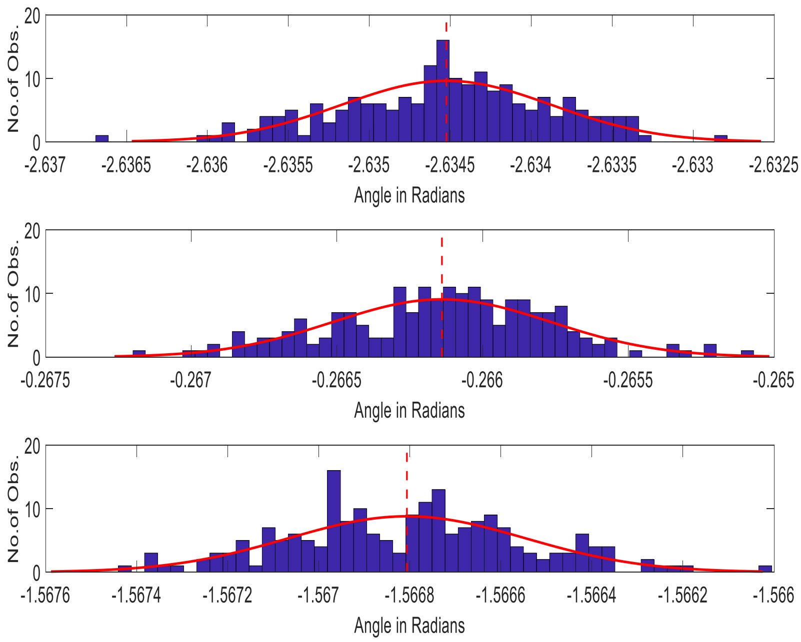

Figure 10: Trailer position left to Van

One observe the resolution of the angle in a range of 0.001°. The precision of the system is judged with the 1-sigma of the distributions of the angles. Additional to the highest available data rate, 27 Hz, averaging on the solution was performed in order to reduce numerical errors and noise. The averaging took up to 5 consecutive measurements, not overlapping. All available standard deviations of the angles are collected in one figure [11].

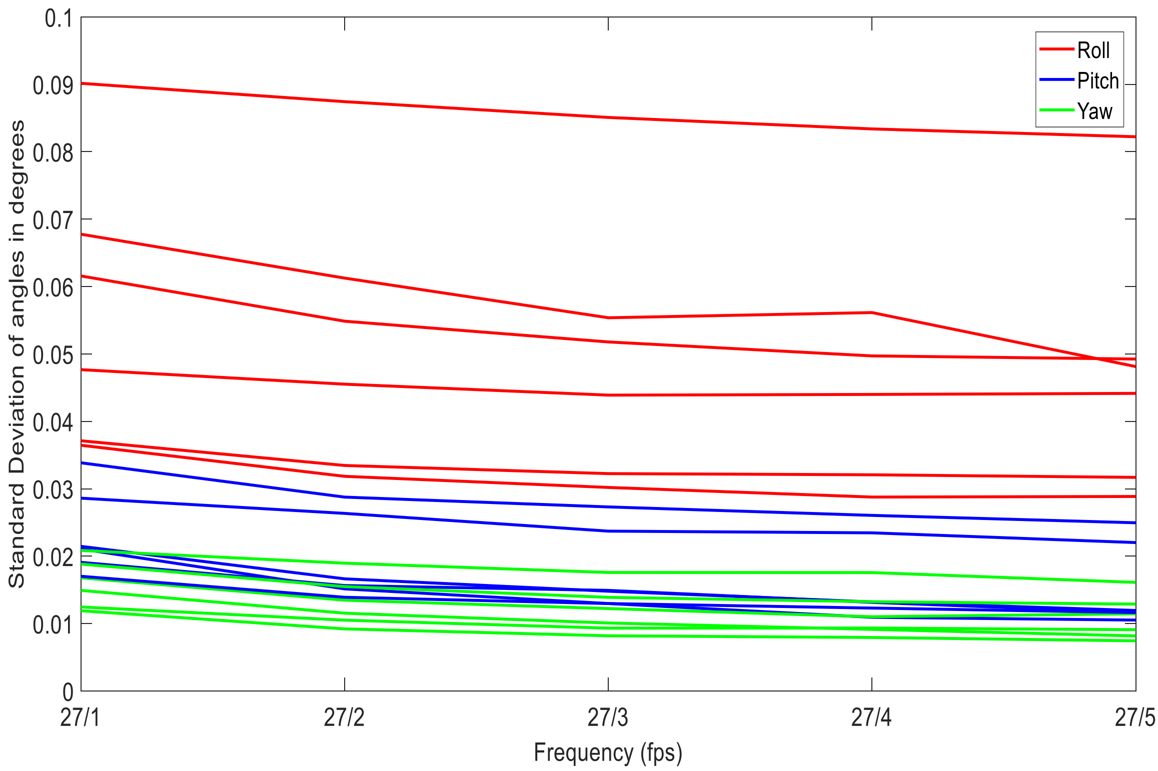

Figure 11: Standard Deviation vs Frequency of frames with different measurements scenario

On that figure on can see that the standard deviation in total does not get bigger than 0.1°. This parameterization describes precision of the vision attitude determination system used in the testing. On the top of this the averaging does the improvement of the system precision only slightly. This enables to use the maximum output data rate without compromising the system precision. The high precision seen in this system is to be addressed, among the others, the uncompressed source data.

It cannot be judged, however, the accuracy of the overall system basing on that output, until this measurement is compared with an external system. That benchmark is performed thanks to simultaneous measurements by the camera system and by the theodolite ranging and bearing. The table represents the averaged attitude values for all the different scenarios with camera measurement and reference measurements. For all the different scenarios, absolute angle difference between two measurements lies under 1.8 °.

	Reference Measurement using Theodolite (Theodolite reference frame, Radians)			Measurement using Camera (camera reference frame, Radians)			Abs. Angle Difference (Degrees)
Scenario	Roll	Pitch	Yaw	Roll	Pitch	Yaw	
#1	-1.399	1.563	1.402	-2.589	0.008	-1.567	**1.696**
#2	1.504	1.293	-1.500	-2.634	-0.266	-1.566	**0.631**
#3	-1.543	1.296	1.540	-2.626	0.279	-1.575	**0.356**
#4	2.425	1.488	-2.634	-2.528	-0.054	-1.774	**1.536**
#5	-2.433	1.480	2.654	-2.529	0.0067	-1.374	**1.742**
#6	-2.968	1.439	2.996	-2.465	0.026	-1.549	**1.413**

Table 4: Absolute Angle Difference for Different scenarios

CONCLUSION

With the executed test it is presented, that the camera system is very precise (0.1°, 1-sigma) and it can deliver 27 Hz rate attitude output. Secondly, the goal of accuracy determination absolute error within the range of 2° is reached and confirmed in the static tests.
One can now only draw a conclusion that the room for the accuracy improvement is still possible to obtain. There are many stages, where it can be lost, for example in accuracy of the vehicles geometry determination or in the fine tuning of the camera parameterization.

ACKNOWLEDGMENTS

This work was conducted within the project NavEn, Funded by the German Federal Ministry for Economic Affairs and Energy (BMWi) and administered by the Project Management Agency for Aeronautic Research of the German Aerospace Center (DLR) in Bonn, Germany (grand no. FKZ: 50NA1509)

REFERENCES

[1] R. C. Gonzalez and R. Woods, Digital image processing.

[2] L. Quan, "Linear n-point camera pose determination," in *IEEE Transactions on pattern analysis and machine intelligence 21.8*, 1999.

[3] Y. Xingfang, "A simple camera calibration method based on sub-pixel corner extraction of the chessboard image," in *Intelligent Computing and Intelligent Systems (ICIS)*, 2010.

[4] "OpenCV," [Online]. Available: http://docs.opencv.org/trunk/d7/d53/tutorial_py_pose.html.

[5] S. Mallick, "LearnopenCV," Big Vision LLC, [Online]. Available: http://www.learnopencv.com/head-pose-estimation-using-opencv-and-dlib/.