Development of a Vision-Based 6D Pose Estimation End Effector for Industrial Manipulators in Lightweight Production Environments

Georg Braun
Center for Lightweight
Production Technology (ZLP)
German Aerospace Center (DLR)
D-86159 Augsburg, Germany
email: georg.braun@dlr.de

Christian Nissler
Institute of Robotics and Mechatronics
German Aerospace Center (DLR)
D-82234 Wessling, Germany
email: christian.nissler@dlr.de

Florian Krebs
Center for Lightweight
Production Technology (ZLP)
German Aerospace Center (DLR)
D-86159 Augsburg, Germany
email: florian.krebs@dlr.de

***Abstract*—Carbon fiber reinforced plastics are playing a key role for aircraft constructions nowadays as well as in the future because of the convenient ratio of strength to weight. Due to the growing requirements of this market, an automation of the production process is necessary. Because of the high unit volumes and accuracy required, the use of robots with increased work accuracy is essential. The German Aerospace Center has developed, in an internal cooperation, an end effector for the camera-based determination of its position and orientation in space. This article deals with the construction and structure of the end effector and first experimental results.**

I. Introduction

In the future, the worldwide demand for new airliners will increase significantly because of the rising number of passengers [1].

High fuel costs and foreseeable upcoming requirements to reduce pollutant emissions are leading to an accelerated replacement of older airplanes. This fact also increases the demand for new airliners.

At the same time, the usage of carbon fiber composites in airplanes will reduce the costs and increase environment comparability because of weight savings and the simultaneously decreasing of fuel consumption.

The production process of these components is very expensive and so the pressure of price, the required number of units per time and the demands on quality will increase in the future.

A. Related Work

These requirements can not be met with today's dominant manual production. An economically-designed production under the above conditions can only be an automated production process. For this, the merging of partial knowledge is necessary as well as the optimization of the complete overall process [2].

The Center of Lightweight Production Technology (ZLP) develops different process chains for the automated production of thermoplastic and thermoset parts. These process chains use industrial robots because of their high degree of flexibility. The robots are used for the manufacturing of parts and also for quality assurance tasks. This type of robot has a high repetition accuracy but a low absolute accuracy [3].

However, industrial-suited processes need high absolute robot accuracies, for example in order to obtain a specific desired fiber angle of preforms in the finished part. This angle has a big influence on material characteristics such as the strength of the component [2]. Because of this, the ZLP evaluates different methods to increase the absolute accuracy of robots under process conditions.

One method we want to address is an enhancement of precision based purely on visual tracking. Many approaches in visual tracking use distinctive point features that are redetected in each frame. One example would be KLT [4]. Instead of using point features, Azad et al. use 3D models by rendering their edges and computing the distance to the corresponding edges in the camera image [5]. Instead of natural landmarks in the camera image, artificial landmark based methods are very common, like for example the ARTag [6] and the AprilTag [7] systems. They proved to be robust and precise features to estimate the position and rotation between camera and marker. They are widely used in mobile robotics [8].

II. Application

The Center of Lightweight Production Technology has developed a process chain for the automated production of large airplane parts with preforms. This is shown in Fig. 1 .

The first step is the automated supply to a roughly defined area and the contour detection of the preform. The position and orientation can hereby be calculated. In the next step, the preform is placed into the mold and fixed. In the third step, the vacuum infiltration of preforms follows with resin and the curing. The last step consists of part machining to obtain the final dimensions and geometry.

The first and the second process steps will be implemented with a robotic portal system (MFZ) (see Fig. 2).

For this process chain, industrial robots are very useful because of their flexibility, the cost advantages for high numbers of units and the constant quality of work results. The illustrated robotic portal system design (*MFZ*) and the two robots on a linear axis (*TEZ*) enable a high level of flexibility but bring a load-dependent deformation of the robots. For the production of airplane parts it is necessary to improve the absolute accuracy of the robot. Because of this, the German Aerospace Center (DLR) evaluates different methods to increase the working accuracy of industrial robots.

The measuring or calculating methods deliver data of position and orientation. This determined data can be used for any

 DOI: 10.1109/ETFA.2015.7301469

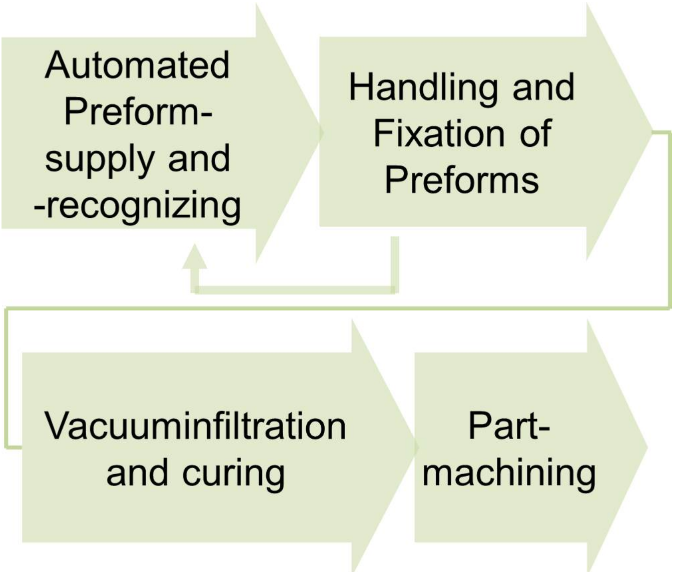


Fig. 1. An example of a typical process chain, showing the steps needed for automated preform-based manufacturing

Fig. 2. Two industrial manipulator systems used at the Center for Lightweight Production Technology : a robotic portal system in the background (*MFZ*) and two robots on a linear axis (*TEZ*) in the foreground

necessary correction of movements.

The first method is the calculation of load dependent deformations. The deformations of robots and the portal system can be calculated only with computational intensive algorithms. This deformation can be calculated only if all the present stresses are known at the same time. This is currently not possible in real time.

The second method is the monitoring and the correction of robot motion with a laser tracker. However, the laser tracker and the needed fiducial reference point attached to the robot are very expensive and the laser tracker needs an unobstructed view of the fiducial reference point. When the tracker loses the line of sight to it, it can no longer calculate the actual position and orientation anymore and therefore this information is not conveyed to the robot. In addition, it can't continue tracking, but has to find the fiducial reference point again.

The third method is the camera-based determination of position and orientation of an end effector in space. The camera at the end effector determines its position and orientation by analysing the images of markers (see Fig. 5). The end effector presented in this publication offers the possibility to determine both the position and the orientation. The aim of the first work is to show the design of this end effector with respect to the given environment and to test the accuracy of the camera-based measurements in a way close to the actual industrial processes.

Especially the first and the second steps of the process chain (see Fig. 1) require a high accuracy of the industrial robots.

Inside the ZLP, a part of these preform process steps can be realized with a robotic portal system (MFZ) (see Fig. 2, robotic portal system in background).

A. *Process Environment*

Attempts to test the end effector were carried out based on the capacity of the TEZ (see Fig. 2, robots on linear axis).

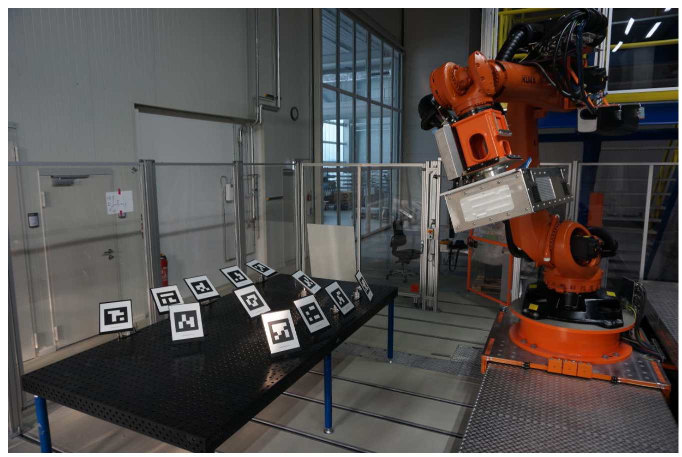


Fig. 3. The process environment, showing the AprilTags markers attached to a table on the left side and the industrial manipulator used (KUKA KR210) with the attached end effector described in section II-B

The test environment is clean and well lit. The robot on the linear axis offers many opportunities for the positioning and orientation of the end effector. The network connection, the power supply and supply of compressed air are performed by the robot equipment. The markers are movable and connected to the table and can be attached variably on the table. The kinematic principle of the attachment is shown in Fig. 5.

The ambient temperature is nearly constant through the use of an air conditioning system. The air contains conductive carbon dust. This dust can cause damage to electrical equipment such as computers.

B. *Design*

The measuring system consists of the end effector and several markers. For the determination of position and orientation, additional markers are necessary (see Fig. 3).

The design of both components is largely determined by the environmental conditions (see chapter II-A) and the development objectives . The overview of the design of the end effector is shown in Fig. 4.

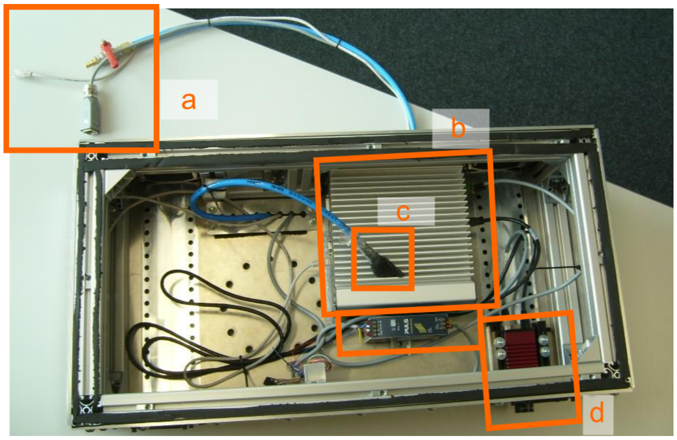


Fig. 4. Design of the end effector, showing the ethernet, air pressure and electricity connections (a), the embedded PC (b), the air pressure cooling line (c) and the camera (d)

The end effector consists of a closed housing which is easily mounted, dismantled and transported manually, using grab bar openings. The housing consists of a solid aluminum frame with numerous mounting options and different designed aluminum sheets. The sheets are screwed with the frame and both parts are connected by black seal strips. These strips seal the parts against each other. The housing is provided with four recording surfaces for a machine control sensor (here a Leica T-Mac sensor). These surfaces serve as a fiducial reference point, called T-MAC, for the laser tracker. By using these fiducial points and the sensor later on, comparative measurements can be performed. This serves as ground truth data which allows an accuracy estimation of the camera-based measurements.

The housing can accommodate an evaluation computer as well as four cameras and the media and energy supply of the internal components. The evaluation computer (see Fig. 4, b) includes an Intel Core i7 quad-core processor and a SSD hard disk and works with the operating system Windows Embedded 7. In addition the computer is shock and vibration resistant and thus very well suited for this application.

The camera is a monochrome camera (see Fig. 4, d) with a 2.3 inch CCD sensor. The resolution is 1600×1200 pixels and the frame rate is maximally 66 frames per second. The camera has a Gigabit Ethernet interface. The attached lens has a focal length of 35mm. The camera mounts can be exchanged. This makes it easy to install different cameras.

The housing has two openings at each long side for the lens of the camera and two openings at each long side for air exchange (see Fig. 4).

Within the housing, the cameras can be moved along their longitudinal axes. This makes it possible to use different lens lengths while keeping the outside dimensions constant. It also serves as a protection for the sensitive camera electronics and optics against collisions by keeping protuding elements at a minimum. The intrusion of electrically conductive dust is prevented by an air overpressure. The air flows out via the camera openings and through the intended air exchange openings. The resulting air flow prevents the intrusion of harmful carbon fiber dust from the outside and at the same time cools the evaluation computer and the cameras (see Fig. 4, a and c).

Fig. 5. The mounting of a marker, showing the AprilTag on the left and the adjustable attachment of the marker to the base on the right side

III. Experiments

In order to make a statement about the achievable precision of the pose estimation system, a series of realistic experiments were performed. The setup used consists of a KUKA KR 210 robot with the developed end effector explained in II-B attached and a table with 10 affixed AprilTags markers [7]. Three different motion profiles of the end effector were executed, two semicircles horizontally and vertically around the markers and a linear movement towards the markers, starting at the most distant position of the robot's working space at approximately 4 meters distance and moving towards approximately 50 cm distance. In Fig. 3 we show the described experiment setup, showing the table with the AprilTags on the left and the robot used with the attached end effector on the right.

In order to obtain ground truth, a laser tracker (Leica AT901) was used, which provides a translational accuracy of 0.5 μm and an angular precision of 0.14 arc sec [9]. Then the motion profiles were executed and to certain time steps the movement stopped and the position and rotation based on the camera based pose estimation and based on the laser tracker were saved and compared. The laser tracker does not however track the same frame of reference as the camera based pose estimation system, but a fiducial marker affixed to the end effector, the so-called

Leica T-MAC. This means that to compare the obtained data, the rotation and translation between the camera and the T-MAC have to be known.

A. Calibration Process

In the shown manufacturing environment, it is important to obtain an absolute positioning and orientation, meaning in reference to a predetermined, common (world) coordinate system. In order to do this, the position of the AprilTags and their size has to be known beforehand. This process is conducted prior to the actual experiments and is called the *marker calibration* step. Hereby the location and orientation of each marker is measured precisely by the laser tracker (Leica AT901) mentioned before. The size of every marker, meaning the size of the outer black border, see e.g. Fig. 5, is measured precisely by a high-precision microscope.

In order to obtain reliable and precise camera based measurements, the intrinsics of the camera, i.e. the focal length and the principal point, have to be known. As mentioned before, also the rotation and translation between the camera frame of reference and the T-MAC have to be calculated. This problem is called a *hand-eye calibration* problem [10], [11], where an unknown pose of a camera and its intrinsics are obtained in parallel. This is conducted by taking many pictures of a precisely known calibration pattern, in our case a chessboard style pattern, at different poses of the end effector. The calibration software then finds the optimal solution for both intrinsics of the camera and the sought transformation between camera and laser tracker frame of reference. The software we used was the DLR Calde and Callab calibration software [12].

B. Image Processing

We used a C++ port of the original AprilTags algorithm presented in [7]. We chose the AprilTags because it improves upon previous ones like the ARTags [6] by offering a higher precision and better robustness [7]. In order to compare the estimated poses, rotations and translations based on the camera measurements (i.e. the output of the pose estimation algorithm) are saved in the very step along with the corresponding marker id as well as the ground truth, which is based on the laser tracker measurements.

C. Experiment Evaluation

Without loss of generality, we choose the laser tracker frame of reference as our common world coordinate system. We therefore want to evaluate the transformation $\mathbf{T}_{lc}$, which is the transformation from the laser tracker to the camera frame based on the pose estimation of the camera.

As ground truth, we compare with the transformation $\tilde{\mathbf{T}}_{lc}$ from the laser tracker to the camera frame, which is based on laser tracker measurements.

Because we want to evaluate the absolute error between our camera based pose estimation and the assumed ground truth we compare the measured transformation based on the pose estimation to the ground truth transformation:

$$\mathbf{E} = \mathbf{T}_{lc} * \tilde{\mathbf{T}}_{lc}^{-1} \tag{1}$$

which is our error matrix. We can split this error matrix E in its translational vector t_E and its rotational part described by the axis-angle representation $(\mathbf{v}, \alpha)$. Then we can evaluate the translational error and angular error:

$$e_t = \|\mathbf{t}_E\|_2 \qquad e_r = |\alpha| \tag{2}$$

D. Experimental Results

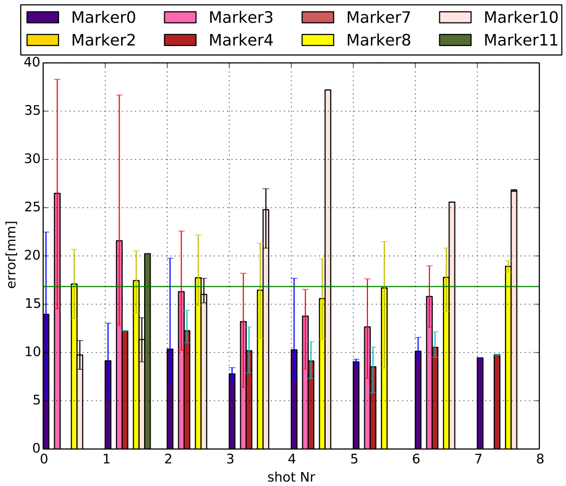
Fig. 6 shows in the first column the translational errors (e_t in equation 2) and in the second column the rotational errors (e_r in equation 2) of the motion profiles explained before, namely a horizotal semicircle movement (first row), a vertical semicircle (second row) and a linear motion towards the table (third row). The individual mean error fluctuates from 4mm to 37mm, but averages to about 15mm in all measurement rows. As expected, the error significantly increases with increasing distance between camera and marker.

One interesting characteristic can be seen in the linear motion in Fig. 6 (e): the error decreases with decreasing distance to the marker (the robot approaches the marker from left to right), but after reaching a minimum of about 5mm it increases again. This can be explained by the optics of the camera: if we get too close the image gets out of focus and therefore gets blurry, making the measurement imprecise. During the experiments it was shown that it is very hard to find a setting for the camera, in which the whole working space of the robot is covered optimally. On the other hand it is possible to choose the camera settings to fit the needs of the application: We choose to try to cover most of the working space; it would however also be possible to set it up in a way to have the best precision at a close while and sacrificing precision when further away.

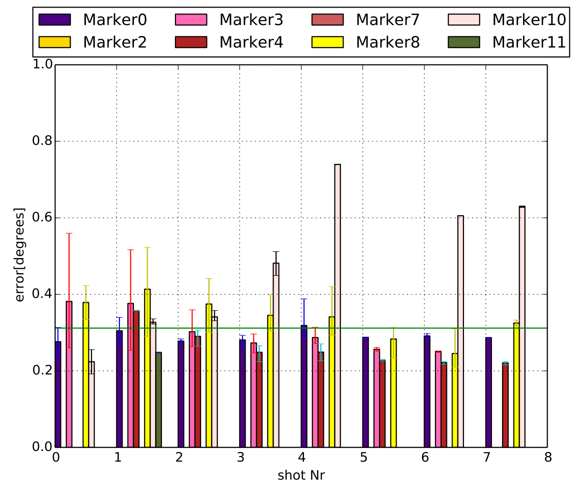
IV. Conclusions and Future Work

We presented a newly developed end effector to estimate its global position and rotation of it to a world coordinate system, based purely on visual measurements. The end effector is specifically tailored to be used in the environment of lightweight production, specifically the handling of carbon fiber composites. We showed the achievable precision by using only visual information of the camera-marker system in a series of measurements with motion profiles most relevant to the application scenario. We thereby demonstrated that by mounting a camera to the end effector of a industrial manipulator no expensive additional hardware (like e.g. a laser tracker) is needed for a precise positioning. However, an extensive calibration step is essential.

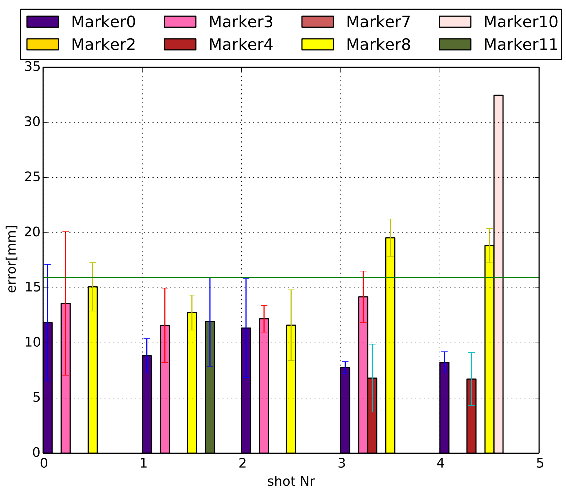
We plan to extend this work by incorporating dynamic measurements. In the shown experiments only static measurements, meaning when the end effector was standing still, were presented. Our setup however also allows for dynamic measurements, only limited to the frame rate of the cameras. This would allow to make statements about the precision of the pose estimation more close to reality in lightweight production environments. This brings up the need for a precise



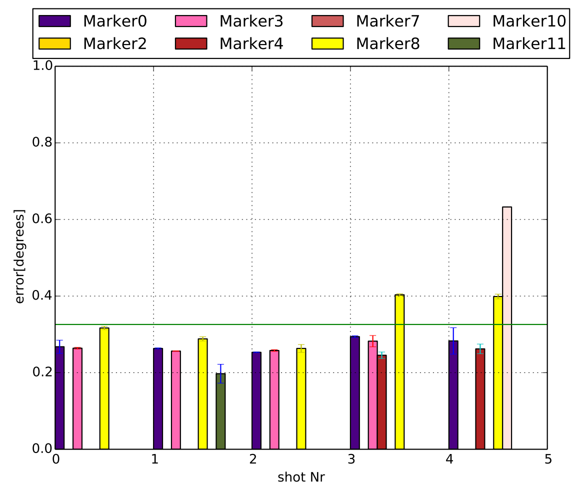
(a)



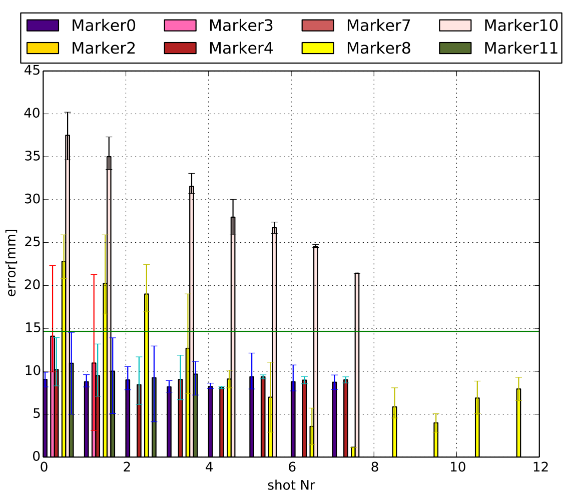
(b)



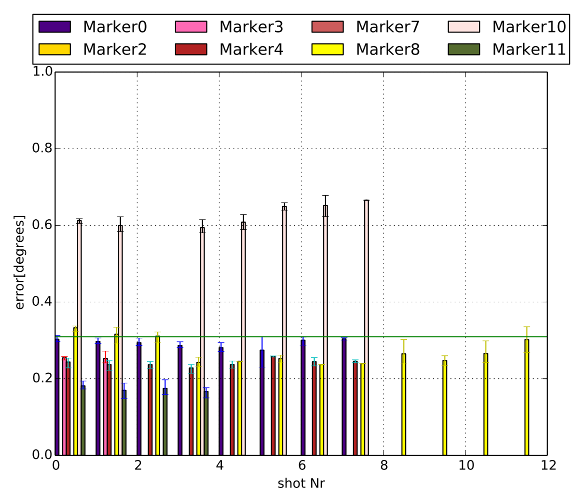
(c)



(d)



(e)



(f)

Fig. 6. mean of the translational (first column) and the rotational error (second column) of the pose estimation based on individual AprilTag detections in mm/degrees, for
(a), (b) : horizontal semicircle
(c), (d) : vertical semicircle
(e), (f) : linear motion.
Note that the error bars depict the minimal and maximal error and the blue line is the total average over all measurements.

time synchronisation of the laser tracker ground truth and the camera information, which is challenging.

A further line of research is to enhance the precision by computer-vision based optimization methods like e.g. RANSAC methods [4] and/or incorporating additional sensors.

References

[1] Airbus, "Airbus global market forecast 2010 2029," Airbus Group, Tech. Rep., 2010.

[2] Dirk Biermann and Werner Hufenbach and Guenther Seliger, "Serientaugliche Bearbeitung und Handhabung moderner faserverstaerkter Hochleistungswerkstoffe," Technische Universitaet Dresden - Institut fuer Leichtbau und Kunststofftechnik, Tech. Rep., 2008.

[3] F. Krebs and S. Nuschele, "Roboter prueft Qualitaet waehrend der Produktion," *Maschinenmarkt*, 2012.

[4] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[5] P. Azad, D. Munch, T. Asfour, and R. Dillmann, "6-DOF Model-based Tracking of Arbitrarily Shaped 3D Objects," 2011.

[6] M. Fiala, "Artag, a fiducial marker system using digital techniques," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2. IEEE, 2005, pp. 590–596.

[7] E. Olson, "Apriltag: A robust and flexible visual fiducial system," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 3400–3407.

[8] E. Olson, J. Strom, R. Goeddel, R. Morton, P. Ranganathan, and A. Richardson, "Exploration and mapping with autonomous robot teams," *Communications of the ACM*, vol. 56, no. 3, pp. 62–70, 2013.

[9] L. Geosystems, "PCMM system specification," Tech. Rep., 2013.

[10] K. H. Strobl and G. Hirzinger, "Optimal hand-eye calibration," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*. IEEE, 2006, pp. 4647–4653.

[11] K. Strobl and G. Hirzinger, "More accurate camera and hand-eye calibrations with unknown grid pattern dimensions," in *IEEE International Conference on Robotics and Automation, 2008. ICRA 2008.*, May 2008, pp. 1398–1405.

[12] K. H. Strobl, W. Sepp, S. Fuchs, C. Paredes, and K. Arbter. (2010, July) DLR CalDe and DLR CalLab. Institute of Robotics and Mechatronics, German Aerospace Center (DLR). Oberpfaffenhofen, Germany. [Online]. Available: http://www.robotic.dlr.de/callab/