IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 12, 2020, accepted November 20, 2020, date of publication December 11, 2020,
date of current version December 28, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3043414

Stereovision Tracking System for Monitoring Loader Crane Tip Position

MAREK GRUDZIŃSKI, ŁUKASZ MARCHEWKA, MIROSŁAW PAJOR, AND RYSZARD ZIĘTEK
Faculty of Mechanical Engineering and Mechatronics, West Pomeranian University of Technology, 70310 Szczecin, Poland

Corresponding author: Łukasz Marchewka (lukasz.marchewka@zut.edu.pl)

This work was supported in part by the project under Grant RPZP.01.03.00-32-0004/17, in part by the European Union from the European Regional Development Fund under the Regional Operational Parogram of the West Pomeranian Voivodeship 2014-2020, in part by the Ministry of Science and Higher Education, and in part by the NCBiR project, ''The use of augmented reality, interactive voice systems and operator interface to control a crane,'' under Grant PBS3/A6/28/2015.

ABSTRACT This paper presents the idea and experimental tests of a non-stationary stereovision system used for supervision of a loading crane. The system based on two independent gimbals, equipped with cameras, allows simultaneously to measure and track a group of individual spatial points. The four steps calibration procedure is required for each gimbal axis, camera optics, gimbals positions, and loading crane position. Modern camera systems have been widely applied in several fields of the industry allowing for accurate measurements and real-time quality control. The systems based on stereo camera models and appropriate image analysis algorithms can measure a group of spatial points, complex surfaces, small displacements, or vibrations. There are also several vision systems such as laser tracker, camera mounted on fully controlled rotating head (called gimbal), or TOF camera, which can be used to follow and measure real-time object displacements. Unfortunately, they all have significant individual limitations explained in this paper. Accuracy tests of 3D measurements including active tracking of length pattern and crane tip have been performed within a selected working area. Based on reconstructed 3D data points, different types of geometric errors have been presented. The results revealed effects of accuracy degradation on the workspace boundaries and for higher observation angles, as well as 3D data bias suggesting mechanical issues. Several improvements in calibration and gimbal construction might result in far enough accuracy for normal operations of the crane.

INDEX TERMS Mechatronics, cameras, calibration, machine vision, stereo vision, remote monitoring, computerized monitoring.

I. INTRODUCTION

A. AIMS AND MOTIVATION

Large development of automation in manufacturing, internet networks, big data processing as well as human-machine interfaces relates to the overall concept of the last industrial revolution, called Internet 4.0 or Industrial Internet of Things [1], [2]. This concept is associated with an increase in the use of sensors, which have the ability to collect data and enable the machines to exchange information and control each other. One component of the Industry 4.0 is broadly understood image data collection and processing, commonly found under the term ''Machine vision''. A high volume of 2D and 3D data recorded by vision equipment could be used to identify shapes, coded characters or defects, and enable fast real-time monitoring in smart factories.

Optical non-contact measurement techniques are one of the most advanced solutions for industrial automation, increasing the efficiency of manufacturing and the final quality of a product. Depending on the application, optical measurements could be divided into two different groups: two-dimensional and three-dimensional. Specific devices are those equipped with two or more integrated cameras, allow to observe a workspace simultaneously from different positions [3]. Thus, stereoscopic spatial observations are possible.

Most of the transport machines, such as loading crane, are controlled manually using hydraulic control levers or, in more advanced solutions, by a remote control panel. Both solutions require the operator's great spatial imagination and eye-to-hand coordination. This paper presents a non-stationary

The associate editor coordinating the review of this manuscript and approving it for publication was Ning Sun.

stereovision system, which is based on traditional cameras installed on two controlled and entirely independent gimbal heads, separated from crane construction. This system allows an operator to observe and actively follow the tip of the loading crane in a specified workspace. In this research, the crane Hiab XS 111 (Cargotec Corporation, Helsinki, Finland) was used. Images captured by the cameras are used as the element of feedback control for servo drives of each gimbal axis. The proposed solution, based on camera models, camera calibrations, and image analysis, is capable of tracking and reconstructing 3D coordinates of a hanging load at the same time and does not require centering the tracked object in the images. Thus, the system accepts even longer delays in object tracking and can operate smoother, which still does not affect 3D reconstruction.

The aim of this paper is to assess the accuracy of the described non-stationary stereovision system, by executing a series of rotations of the crane around its main axes. Captured and reconstructed positions of the crane tip were compared with a model of the boundary workspace. Experimental measurements must have been preceded by complex calibration of gimbals kinematics, cameras optics and positions, and the crane position in the common workspace. In the last chapter, sources of potential measuring errors and possibilities of future evaluation of the proposed system are discussed.

B. RELATED WORKS

The problem of 3D reconstruction of a scene in stereo camera systems can be solved using several techniques, which commonly base on image differences detection and analysis for selected regions of interests. For less demanding real-time applications the image disparity can be estimated by combining the SGM algorithm, local similarity metrics, and by analyzing the occlusion effects [4]–[6]. The major disadvantage of this technique is poor performance in the 3D scenes with a low level of ambiguity. i.e., large uniformly colored surfaces. Another problem appears in the environment consisting of numerous similar features that are difficult to distinguish and match in the stereo images. A more accurate global approach tends to come at a high computational cost and cannot be used for real-time reconstruction [7]. In several studies modifications of SGM algorithms have been proposed which provided shorter running time of main application. The proposed in [8] method of stereo block matching (SBM) achieves running time of 0.3 s per pair of stereo images and cannot be applied in real-time tracking.

Another approach to reconstruct 3D scene is the use of the Structure from Motion (SfM) algorithm, which is fundamental for the photogrammetric systems or aerial photography in geotechnics. The algorithm computes 3D points (structure) as well as the camera intrinsic and extrinsic orientations (motion) for a set of images by using the bundle adjustment algorithm that minimizes the so-called points reprojection errors [9]. This method is more efficient for highly detailed scenes containing various shapes, edges, and colors, but can produce errors in uniform areas (i.e., grass, trees, sand) [10]. The local similarities and corresponding points need to be first identified in the set of images which is a time-consuming process. Although the SfM can produce a high accuracy point cloud for correspondent image features (1-10 cm in aerial photography [11], [12]), the remaining points can produce a large amount of noises or surface distortions.

In more demanding applications the problem of ambiguities can be solved by marking the scene using characteristic reference points or by projecting light patterns. Further calculations can use the so-called inverse camera projection models with triangulation techniques and elements of epipolar geometry to reconstruct 3D points, even for each image pixel [13]. The time-consuming computations of image similarities are not required, and close-range systems can achieve an accuracy of 0.01 mm. A relatively small number of reconstruction errors can be detected and eliminated in post-processing of the point cloud. Although the structured light projection allows for fast and accurate measurements, it still can be proceeded only in a small measuring volume.

The most advanced cameras are characterized by high resolutions sensors with a diagonal of 30 mm or more and they are equipped with high-end lenses, which result in low noises and highly detailed images. For this reason, they are commonly used in most accurate 3D scanners and photogrammetric cameras [14]. The latest camera models such as Ximea xiB-64 can achieve up to 70 fps with a resolution of 65 MPix and can be utilized. for real-time object tracking. However, the processing of such a large amount of data requires robust hardware components and efficient software computations. Moreover, these cameras are relatively expensive and the measuring volume is still limited by the camera field of view.

A different approach needs to be used for the purpose of large volume observations. There are various integrated tracking systems usually based on a single fully controlled gimbal head. These trackers can be equipped with a laser light source and allow to measure angles and single distance, or with a camera sensor allowing only for angular measurements. The first of them can be used for quality control of large-scale construction such as airplane wings, while the second one can be applied for target tracking in e.g., in urban monitoring or military drones [15]–[17]. Nonetheless, the single camera does not allow for distance measurement in a simple manner. Here, the only approach is the use of the photogrammetric method where a single camera must be displaced to over a dozen positions around a static object in order to capture images. Then, the problem of points correspondence, camera calibration, and points reconstruction can be solved in one optimization process [18]. For fast 3D measurements, two or more cameras located in specified positions are necessary. The integrated stereo-pair camera system could be utilized, however, in order to measure a large-scale objects, the cameras need to be placed in a long relative distance.

The Lidar system, by measuring the time of flight of a rotating laser beam, can scan the 3D geometry of a large workspace and follow separated objects at the same

time [19]–[22]. Apart from very expensive devices, this method provides a relatively low angular resolution of data points and makes difficulties in the recognition and distinction of small objects. Moreover, Lidar systems in principles are unable to extract image features such as shapes, colors, or edges, which is essential in high-end measuring systems. An advanced algorithm needs to be applied to extract and recognize objects in a large point cloud. The problem can be solved by applying additional cameras and proper calibration procedures [23]. Another new solution for industrial applications are the time-of-flight cameras (TOF), called also depth-cameras. These devices installed on a gimbal head allow to follow and measure 3D coordinates of the visible scene using an integrated IR illuminator and single IR sensor. However, they have significant limitations i.e., low sensors resolution, sensitivity to the sources of external strong IR radiation, and low accuracy of spatial and angular measurements [24].

In order to simplify the crane control, several modern solutions were recently invented i.a. telemanipulation or teleoperative systems, which are based on force feedback [25], [26]. In [27] the authors proposed the six-axis tensometric handgrip, directly attached to the crane tip, which allows the operator to move the crane in desired directions using force input of the operator's hand. Another telemanipulation solution may be an external exoskeleton for the operator, which facilitates the control of a load position through simple and intuitive boom movements. However, all these solutions do not give any information about the actual exact position of the crane tip [28].

Assessment of position or displacement of a load in the long-range observation causes many difficulties to the operator or even remains impossible. Theoretically, the crane could be equipped with additional distance indicators or angular encoders, allowing to calculate the tip position. However, such solutions do not include significant deflection of the crane construction under the full load or load fluctuations. Authors of a paper [29], [30] tested a single camera system to control the crane position by using markers attached to the operator's hand. In the paper [31] authors proposed a laser tracker integrated with a single CCD camera, which observes and predicts the hook position in real-time, reducing the problem of image processing delay. However, the device performs simple distance measurements, without system calibration in relation to the crane coordinate system. The authors extended their research in [32] utilizing three markers attached to the crane base in order to determine only the main rotation axis of the crane. In another research, the human-machine speech interface was proposed to interact and control the crane movements by the use of natural language [33]. Although these systems utilize built-in encoders, they are still designed to initialize and sustain the movements and not to measure the actual crane position.

Several optical approaches to control the load with the use of fuzzy logic, neural networks, and anti-sway algorithms have been described [34]. However, the authors of this review paper did not present their own working solution and did not provide accuracy tests of discussed stereovision systems. The applied vision systems were based on fixed cameras, which might limit their workspace.

Camera systems have been also utilized for observations of a hanging load. In [35], two cameras have been mounted directly to the crane base on the common beam, which allows for simultaneous vertical rotation. Only the horizontal rotation axes have been designed as an independent. Thus, the system was not able to observe the crane rotation in its base. The authors of the paper assumed that the observed hanging load must be always hold in the centers of the cameras images and thus simplified calculations of 3D coordinates to a triangulation problem. However, measuring accuracy depends strongly on well-established cameras models, images distortions, and homogenous transformations (called intrinsic and extrinsic parameters), commonly used in 3D vision systems. The Authors did not mention about calibration of the cameras positions using calibration patterns and other target points.

At the present time, there are many optical, stationary systems allowing for three-dimensional reconstruction of the position or displacements, e.g., OptiTrack (NaturalPoint, Inc., Corvallis, USA), Vicon (Vicon Motion Systems Ltd, Oxford, UK) or ART (Advanced Realtime Tracking GmbH, Weilheim in Oberbayern, Germany). Stationarity in principle means that all the optical parameters must remain constant during any operations, i.a., camera position and orientation, focal length. The advantage of these ready-to-use systems is a high accuracy (0.1 mm in the scanning area of 9×9 m for OptiTrack Prime^{x} 41 and 0.15 mm for Vicon T40S) [36], [37]. The disadvantage may be measuring volume limitation, due to fixed cameras positions (in relation to the global coordinate system). These systems could theoretically be used to reconstruct the three-dimensional position of the crane, but this would require a sufficient number of cameras covering the desired measurement space.

C. PAPER CONTRIBUTION

This paper presents the original approach to measure 3D displacements of the loading crane using two independent and moveable cameras. The 3D coordinates of a group of spatial points attached to the crane are obtained based on typical camera models and camera extrinsic parameters. However, this solution is more flexible and allows to recalculate the extrinsic parameters continuously by considering the kinematics structure of the gimbals and individual rotation of the camera. In contrary to related works and described techniques, the proposed solution does not require fast tracking algorithms and accepts even large delays in object observations. The presented system is also more flexible in terms of installation and alignment to specific machine construction using separated and independent gimbals. Moreover, this solution should be relatively cheaper and at the same time might generate correct measurements of the crane tip even in wide observation angles.

II. GENERAL IDEA AND CALIBRATION OF NON-STATIONARY STEREOVISION SYSTEM

A standard stereovision system consists of at least two calibrated cameras located in well-defined relative distance and orientations, allowing for three-dimensional measurements in a specified workspace. Such an effective solution allows to take 3D data of immobile or moving objects, as long as the relations of both cameras remain constant and the object stays visible in both images simultaneously [38]. The authors of this paper call it the "stationary system" for further considerations.

The use of cameras in the stereo system requires proper estimation of the field of view and measuring volume. In stationary systems, such as PONTOS Live (GOM GmbH, Braunschweig, Germany), objects can be observed only in a certain limited area in a relatively short distance of ca. 0.5-2.0 m [39]. Measuring volume could be extended using an additional robotic arm and continuous recalculation of the cameras' positions, which could be difficult to adapt for long-range measurements. Low accuracy of manipulator causes inaccurate measurements because an error on the long manipulator arm drastically affects measurement accuracy [40]. Moreover, the difficulty results from the impossibility of interfering with the closed software architecture of the PONTOS system (GOM GmbH, Braunschweig, Germany) when recalculating camera positions. Without using the robotic arm, the distance between cameras needs to be highly increased, but the angular range of measurements remains similar. Theoretically, for greater measuring volume the lenses with shorter focal length (and thus wider observation angle) could be used, however, their noticeably lower optical resolution, high distortions, and image blur on boundaries could disturb details recognition. In this case, even a high-resolution camera could not improve image quality. Such cameras have mostly low frame rate (per second) and are more expensive.

Field of view and high costs may not be the limitations in the case of independently moving cameras, actively following the observed object (Figure 1). A properly adjusted tracking algorithm should only keep the object inside the field of view since the 3D reconstruction can be performed for the whole image. Even small delays in tracking are not a major problem, as long as the tracked object is visible in the camera field of view. Thus, the standard lenses with low distortions as well as inexpensive cameras with low resolution and high frame rate can be used.

In contrary to stationary systems, the actual relative positions and orientations of both cameras in the world coordinate system should be calculated continuously. It faces the problem of continuous determination of parameters associated with the camera and gimbal coordinate systems. For this purpose, complex calibration is necessary, which involves optic parameters, encoders signals, and gimbals kinematics. It is worth mentioning that the camera coordinate system, associated with the main optical axis and sensor orientation, could not be directly indicated by the use of standard measuring

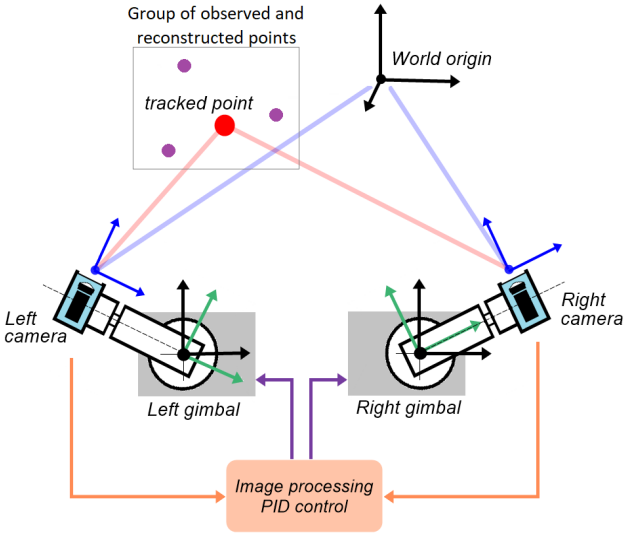


FIGURE 1. The idea of a stereovision tracking system with two independent gimbals. The cameras coordinate systems (blue) rotate independently in calibrated gimbals coordinate systems (green) in order to follow a group of 3D points at the same time. The involved PID algorithms control servo drives based on images feedbacks.

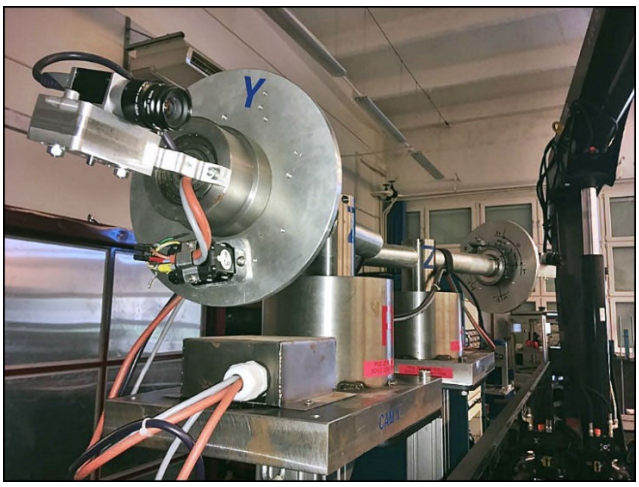


FIGURE 2. A pair of gimbals designed and used for loading crane tracking. The construction was made mainly of steel and aluminum and assembled using large ball bearings.

devices. The so-called camera point, where theoretically the camera is placed, does not relate to the camera housing or the lens mounting.

In this research, the developed stereovision system consists of two industrial cameras scA1600-14gm (Basler AG, Ahrensburg, Germany) mounted on two-axis gimbals, placed on rigid mobile platforms (Figure 2). Rotations around each axis are carried out using Dynamixel MX-64 and MX-28 servo drives (Robotis, Lake Forest, USA), equipped with internal angular encoders. The angular resolutions of the rotations were extended by applying additional reduction gears. The main application for controlling the system was developed in Visual Studio, including OpenCV 3.4.8 (Open Source Computer Vision Library, Palo Alto, USA),

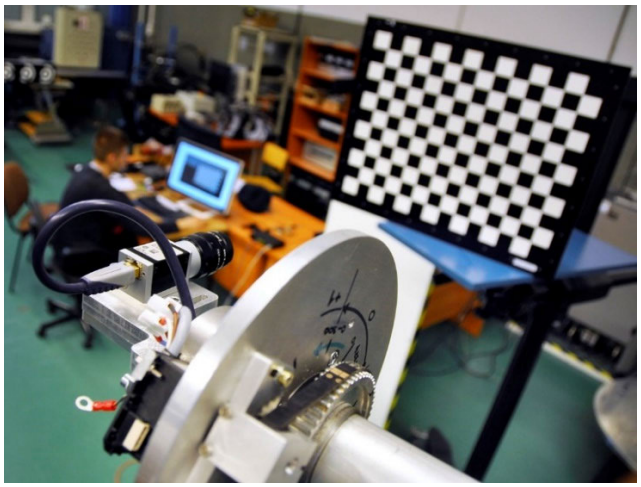


FIGURE 3. Calibration of camera intrinsic parameters using rectangular flat chessboard.

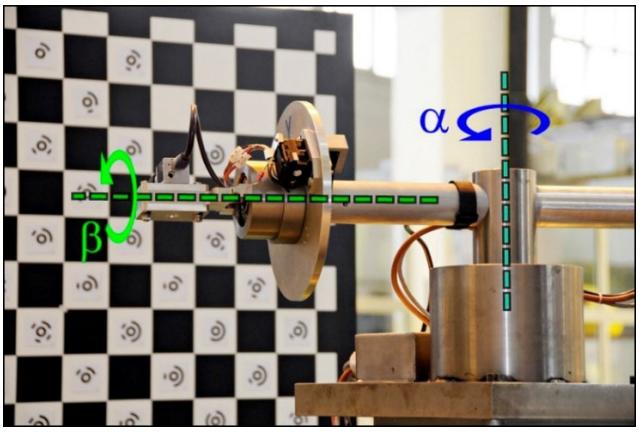


FIGURE 4. Calibration of gimbal axes directions in relation to the camera origin, using large calibration pattern with coded points and by performing two individual rotations around each gimbal axis.

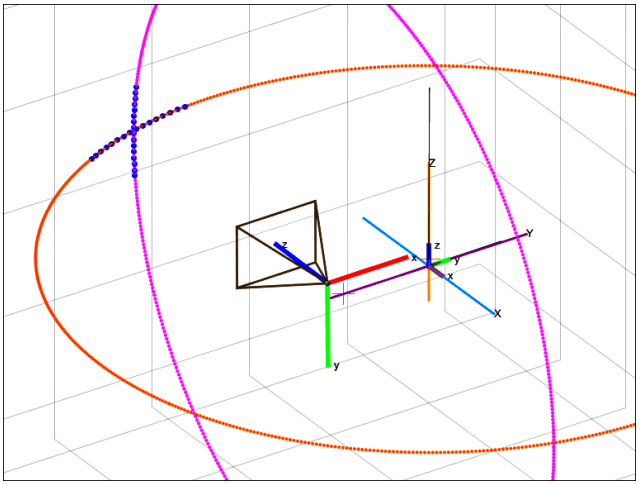


FIGURE 5. Determination of gimbal origin in relation to camera origin. The blue dots indicate the origins of the calibration pattern inside the camera coordinate system. Two circular paths indicate apparent motion of calibration pattern origin in the camera coordinates accordingly to gimbal axes direction. The obtained inverse transformation of both coordinate systems and circular paths were visualized.

Basler Pylon 6.0.1 (Basler AG, Ahrensburg, Germany), Dynamixel (Robotis, Lake Forest, USA), and DirectX SDK June 2010 (Microsoft Corporation, Redmond, USA) libraries. The implemented PID algorithm for active camera tracking was described in another authors' publication [41].

A. CALIBRATION OF THE STEREO TRACKING SYSTEM

Complex calibration procedures for the non-stationary system, which allow for continuous observation and simultaneous 3D reconstruction of the crane tip, were also developed. In the system, a single set of parameters is insufficient due to the fact, that cameras rotate and change positions continuously. Thus, the positions and orientations of the cameras in a common space need to be continuously recalculated. The proposed calibration has been divided into four steps, for which different calibration patterns were designed. Three steps had been explained in detail in another authors' publication [41].

Step 1: Calibration of camera parameters, that describe the relation between the object space (three-dimensional) and its projection on the image plane (flat surface of the camera sensor) [18], [42]. These parameters, called "intrinsic parameters", can be achieved once a time outside the workspace (Figure 3) For intrinsic parameters a regular flat chessboard and several functions from Camera Calibration Toolbox for Matlab (The MathWorks, Inc., Natick, USA) were utilized [43].

Step 2: Calibration of gimbals heads, which determines the initial relation of the obtained camera origin to gimbal base, as well as gimbal position in a common workspace. The presented method is based fully on optical techniques, where gimbal calibration was based on multiple observations of the large calibration pattern separately for the Z and Y axes (Figure 4) and continuous reading of the encoder signal. Both the camera rotations can be transformed into circular motion paths of the pattern around the camera (Figure 5).

Step 3: Calibration of gimbals extrinsic parameters, that determine the positions and orientations of both gimbals origins in relation to the common world coordinate system. This step needs to be performed once a time after mounting both gimbals in the common workspace. Due to the large measuring volume, the calibration pattern was constructed in the shape of a big regular 3D cross with coded markers (Figure 6). The extrinsic parameters were obtained using the Bouguet approach [43]. Calibration of gimbals extrinsic parameters is based on the well-established initial relation of each camera origin to each gimbal base (from Step 2). Since these relationships can be further recalculated any time after reading actual encoders signals, the calibration of gimbals extrinsic parameters requires only new extrinsic parameters for both the cameras in relation to the 3D cross. Then, an inverse task can be performed for any gimbals rotations and camera extrinsic parameters can be continuously recalculated, which is essential for 3D points reconstruction in real-time.

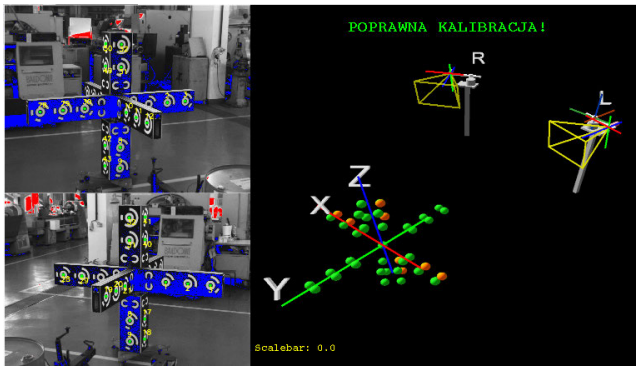


FIGURE 6. Extrinsic parameters calibration for stereo cameras and gimbals. The observed corresponding coded markers were reconstructed (orange spheres) over the reference cross (green spheres).

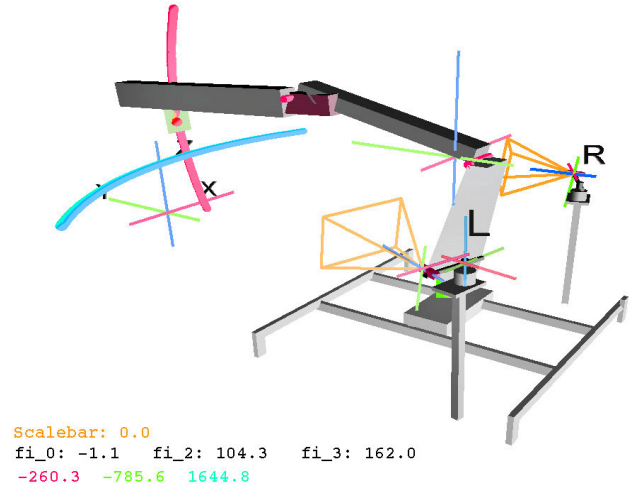


FIGURE 7. Crane position calibration process visualized in the main application window with visible reconstructed trajectories of the tracked marker. At this moment a model of the crane is placed in a previously saved position, for different gimbals positions, therefore its kinematic structure is not correctly recalculated and displayed.

Step 4: Calibration of the loading crane position and orientation. In order to measure the crane tip positions, a relation between the world origin of the stereovision system and the independent coordinate system of the crane must be found. The stereovision system calibrated in the last three steps, capable of tracking and measuring markers, was placed nearby the crane base. A distance of approximately 4 meters between cameras was assumed. Then, two rotations of the crane (around the X-axis and Z-axis) were performed with the simultaneous reconstruction of a marker attached to the tip (Figure 7). For two obtained groups of reconstructed points, two circle models were fitted, which normal vectors indicated two main directions of the crane coordinate system.

As shown in Figure 8, the last section of the crane was marked with a unique flat board with coded points printed on it, easily recognizable by the developed software. The lowest marker on the flat board was assumed to be the tracked marker. The board had been fixed and positioned in relation to the pivot of the hook and direction of the crane boom using the Tritop photogrammetric system (GOM GmbH,

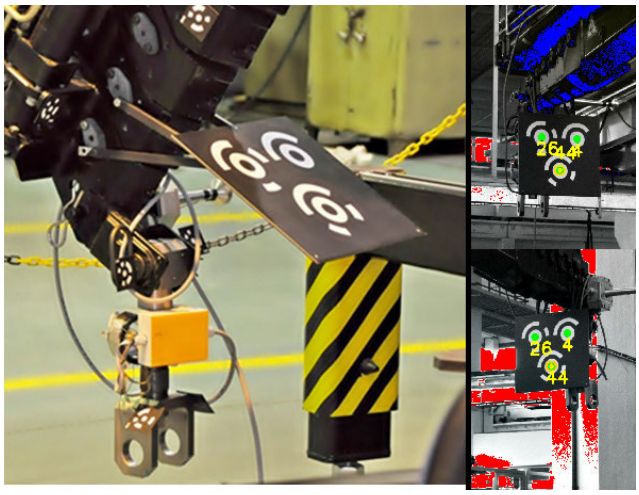


FIGURE 8. Coded pattern indicating the crane tip position, recognized in the pair of stereo images.

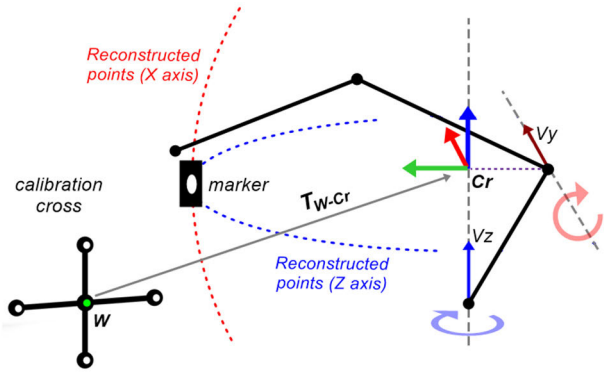


FIGURE 9. The idea of calibration of the crane origin Cr by tracking the marker attached to crane tip and reconstructing its 3D positions in world coordinate system W for two individual rotation of the boom.

Braunschweig, Germany). Since the arrangement of coded markers is well defined on the board, further reconstruction of all the markers allows calculating coordinates of the crane tip. Moreover, the known orientation of the flat board has been used to solve the problem of inverse kinematics and calculate all booms angles. Then, the crane could be visualized in the 3D animation.

According to Figure 9, the position of the crane was considered as a projection of the initial point of vector V_y onto a line, which coincides with vector V_z. Based on the directions of two vectors V_y and V_z the rotation matrix was created, which specified the orientation of the crane in the world coordinate system. Finally, the total transform $T_{W\rightarrow Cr}$ was combined in homogenous form and saved to the configuration file. The computer program was designed to multiply the captured 3D points by the obtained transform automatically and to express the points in the crane coordinates. The described calibration steps allowed performing further experiments.

B. AUTHORS' APPROACH TO 3D RECONSTRUCTION

At least two cameras are necessary for the reconstruction of the spatial coordinates of any points visible in both cameras

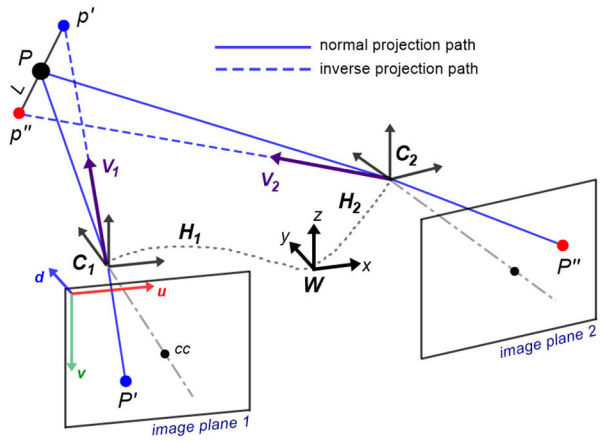


FIGURE 10. The general idea of 3D point reconstruction by performing inverse projection of image point P′ and P″ and finding the closest solution for two lines intersection.

images. In most of the standard applications, the correspondence problem is solved by comparing characteristic image features [44], [45] or using coded targets (markers), which is much more reliable for high accuracy measurements [46].

For the single camera, according to Figure 10, the relation between representative point P and its projection P' on the image plane (camera sensor) can be written as the normal projection equation (1):

$$s_n \cdot \begin{bmatrix} P'_u \\ P'_v \\ P'_d \\ 1 \end{bmatrix}_{4\times n} = \boldsymbol{A}_c \begin{bmatrix} P_x \\ P_y \\ P_z \\ 1 \end{bmatrix}_{4\times n} \tag{1}$$

where:

$$\boldsymbol{A}_c = \begin{bmatrix} f_m & \gamma & cc_m & 0 \\ 0 & f_n & cc_n & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2}$$

is called the camera matrix, which includes several parameters i.e., focal length f_m, f_n, sensor skewness γ, crossing point (q_m, q_n) of the main optical axis, and the sensor. Coordinates of all projected points P' must be normalized with respect to their P'_d coordinate in order to eliminate the unknown scale factor s_n. Thus, all the projected points can represent one flat image plane $d = 1$. For the camera located in a specified position in world coordinate system W, the equation (1) can be extended to:

$$\begin{bmatrix} P'_u/P'_d \\ P'_v/P'_d \\ P'_d/P'_d \\ 1 \end{bmatrix} = \boldsymbol{E} \cdot \boldsymbol{A}_c \cdot \begin{bmatrix} P_x \\ P_y \\ P_z \\ 1 \end{bmatrix}, \tag{3}$$

where $\boldsymbol{E}$ is the homogenous transform of the camera coordinate system C, including rotation and translation. All the necessary parameters of the camera can be obtained after the calibration process [47]. The proposed by authors particular form of camera model allows calculating projection points directly in world origin instead of in the camera coordinate system. Thus, for each CCD sensor, all the 3D points are transformed at first to the individual projection plane d according to the camera matrix $\mathbf{A_c}$ and then transferred within the world coordinate system according to camera extrinsic parameters $\boldsymbol{E}$. The cameras origins are treated as the local coordinate systems in the world coordinate system associated with the calibration pattern. Such an approach allowed to simplify and optimize the program code.

In stereo vision systems, the reverse task is normally performed, where the position of spatial point P is unknown, however, it still can be projected and normalized to the image plane. Then, the projected point P' can be transformed from the image plane to three-dimensional vector v, using the so-called inverse projection model:

$$\begin{bmatrix} v_x \\ v_y \\ 1 \\ 1 \end{bmatrix} = E \cdot \boldsymbol{A}_c{}^{-1} \cdot \begin{bmatrix} P'_u \\ P'_v \\ 1 \\ 1 \end{bmatrix}. \tag{4}$$

Since the scale factor is unknown and the projected point must have been normalized, the length of the vector v does not represent the real distance from the point P to the camera. In this case, the vector v indicates a line on which the point P should lay, which still does not solve the reconstruction problem. Therefore, a second camera localized in a different position must be utilized.

As shown in Figure 10, the point P is observed by two cameras and projected to image planes as a pair of corresponding points P' and P''. By applying the inverse projection model, described in equation 4, both projected points can be transformed to vectors v_1 and v_2 respectively. Assuming known translations t and rotations R of both cameras, two lines in common coordinate system W can be determined, which should intersect in point P. In practice, the lines pass each other by a very small distance L (from p' to p'') as a result of estimated cameras models and calibrations errors. The distance L is the length of vector v_3 obtained from the cross product of v_1 and v_2 and scaled by unknown factor g. Therefore, the reconstruction problem can be simply solved by adding three vectors v_1, v_2, and v_3 in W space and finding their scale factors, using the following equation (5):

$$g_1\vec{v_1} + \vec{t_1} - \left(g_2\vec{v_2} + \vec{t_2}\right) + g_3\vec{v_3} = 0. \tag{5}$$

Finally, it can be assumed that point P is the midpoint between p' and p'' [19] and can be obtained using the following equation (6):

$$\boldsymbol{P} = \frac{1}{2}\left(g_1\vec{v_1} + \vec{t_1} + g_2\vec{v_2} + \vec{t_2}\right). \tag{6}$$

III. EXPERIMENTAL MEASUREMENTS OF STEREO VISION TRACING SYSTEM

The stereo pair of two already calibrated gimbals, placed in the common workspace, could measure the 3D coordinates on condition that the relation of two cameras mounted on

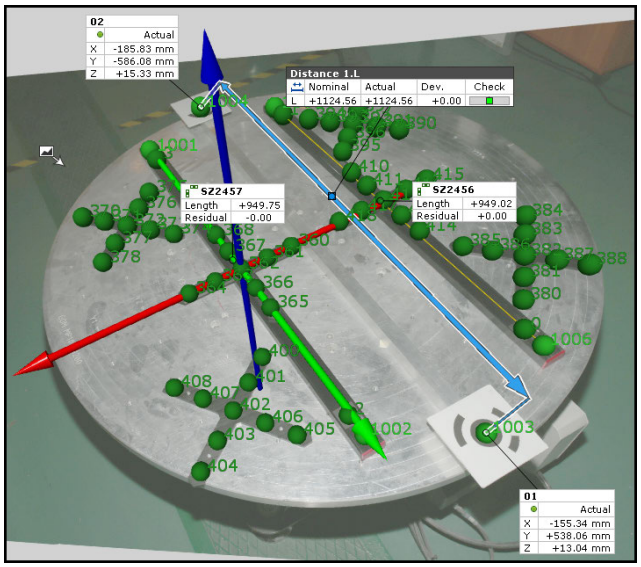


FIGURE 11. Photogrammetric calibration of the aluminum length pattern before system testing, by the use of GOM Tritop system (GOM GmbH, Braunschweig, Germany).

gimbals is defined, similarly to the stationary stereo systems. Although gimbals coordinate systems remain constant, the cameras relative orientation changes together with gimbal axis motions. The extrinsic parameters must be recalculated all the time to provide a reliable 3D reconstruction. In this research, two different experiments have been carried out using flat coded markers, whereby the selected marker was always used to center views of cameras and all the visible markers were reconstructed in 3D space at the same time. For both the experiments the cameras positions and orientations were recalculated continuously based on calibration results and encoders signals, however, in the first test a fixed object in different positions was considered, and the second test was focused on dynamic 3D reconstruction during gimbals movements. Even while static measurements the tracking algorithm remained active.

A. RECONSTRUCTION OF LENGTH PATTERN

The first experiment with a length pattern with two coded markers attached to it was carried out to initially verify the abilities of 3d reconstruction in the non-stationary system. In previous research, the authors of this paper conducted similar measurements using the stationary stereovision system with fixed cameras [48]. In this research, a much larger area was observed by using movable gimbals. The experiment consisted of measuring the length pattern fixed in different positions and with the tracking algorithm enabled. The distance between markers was confirmed by the use of the Tritop photogrammetric system (GOM GmbH, Braunschweig, Germany) and equals 1124.56 mm (Figure 11). An additional middle marker was attached to the length pattern in order to center the cameras views. The distance measured between the markers was displayed in real-time in the main application window (Figure 12).

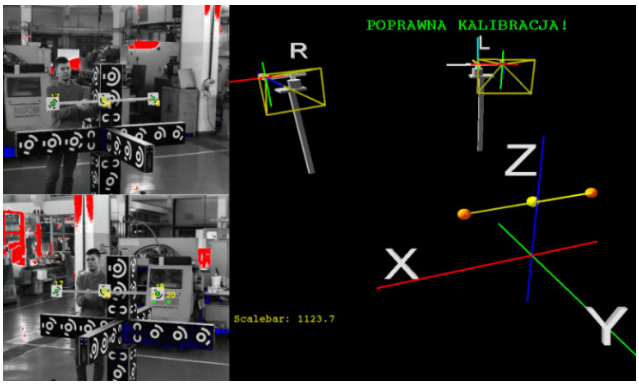


FIGURE 12. The length pattern held in hands and reconstructed nearby the origin of the world coordinate system (calibration cross), where its length was determined most accurately.

TABLE 1. Measurements of the Length pattern (reference: 1124.56mm)

Angle of observation [°]	Result (horizontal) [mm]	Deviation to the reference	Result (vertical) [mm]	Deviation to the reference [mm]
-80	1092.0	-32.6	1095.2	-29.4
-64	1103.2	-21.4	1100.8	-23.8
-48	1112.3	-12.3	1116.7	-7.9
-32	1119.7	-4.9	1120.5	-4.1
-16	1122.3	-2.3	1122.9	-1.7
0	1123.7	-0.9	1124.5	-0.1
16	1127.2	2.6	1126.4	1.8
32	1129.1	4.5	1127.6	3.0
48	1131.2	6.6	1130.1	5.5
64	1136.3	11.7	1135.0	10.4
80	1149.2	24.6	1147.0	22.4

The first measurement was taken nearby the position of the calibration cross (Figure 12). Then, the length pattern was iteratively shifted by a known angle of 16° up to ±80°, around the middle point between two gimbals (assumed to be the center point of the stereo vision system). Moreover, for each point at the path, the measurements were performed in the horizontal and vertical orientation of the length pattern. The results are presented in Table 1. As expected, the results only verified the calibration procedures and overall performance of the non-stationary stereo vision system.

Summarizing, the deviation error increases as the angle of observation increases. In the area between −48° and +32°, the error is less than 10 mm, which seems to be a satisfying result in such a large measuring volume. Moreover, the error increases faster with increasing positive observation angle values which might be a mechanical issue, i.e., gear transmission play or gearbox assembling inaccuracies. The best results, close to the reference value, were obtained near the world origin (position of the calibration cross).

B. DYNAMIC RECONSTRUCTION OF CRANE POSITION

The completely calibrated stereovision system, with established relation to the loading crane, was used to perform the

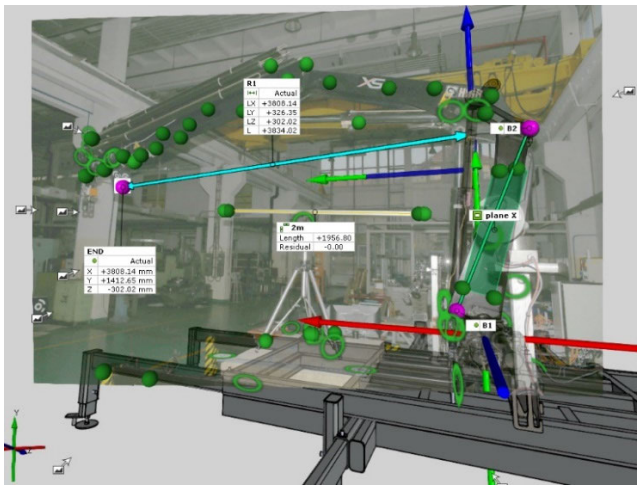


FIGURE 13. The tracked marker (signed as END) and other reference points, aligned to CAD model of the crane Hiab XS 111, measured and visualized using Tritop photogrammetric system (GOM GmbH, Braunschweig, Germany).

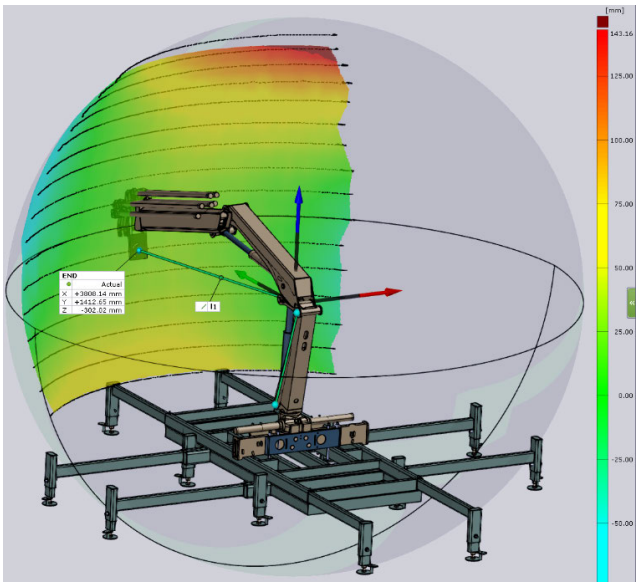


FIGURE 14. Global deviation map of 3D reconstruction errors of the crane tip positions in reference to a spherical model of theoretical workspace visualized in the ATOS Professional software (ATOS Professional V8, Rev. 82961, Build 2015-01-28 by GOM GmbH, Braunschweig, Germany). Points between the recorded trajectories have been interpolated.

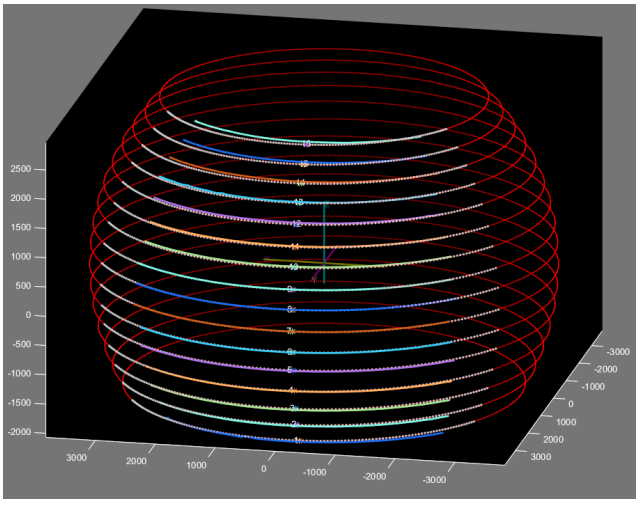


FIGURE 15. Reconstructed motion trajectories of the crane compared with theoretical reference trajectories (red).

main experiment in a large part of the crane workspace. The experiment consisted of camera tracking and simultaneous reconstruction of the crane motions, after applying a constant rotation of the crane around its Z-axis in the range of $\pm 60°$. Each rotation was repeated after lifting boom angle by every 5° around X-axis and exact values of these angles were read with the aid of an accurate protractor. Such an approach allowed further to reconstruct theoretical motion trajectories in Matlab (The MathWorks, Inc., Natick, USA) and compare them to those reconstructed by cameras. The rotations were applied by manual control of the crane. Each measurement allowed to capture ca. three hundred and fifty points in twenty seconds. In total, sixteen representative measurements were taken, covering a large part of the available space, limited only by hall volume and marker visibility in both cameras images.

The experiment was preceded by the photogrammetric measurement of the marker in relation to the base coordinate system of the crane using the Tritop photogrammetric system (GOM GmbH, Braunschweig, Germany). Several characteristic points, e.g., joints centers, were measured in order to align the CAD model of the crane to these points (Figure 13). Thus, an initial angle of the crane boom, as well as the distance of the marker to the Z-axis, have been determined.

The proceeded experiment resulted in sets of reconstructed 3D points which at first have been converted to the mesh structure and visualized using ATOS Professional software (ATOS Professional V8, Rev. 82961, Build 2015-01-28 by GOM GmbH, Braunschweig, Germany), in order to assess the global tendency of measuring error distribution (Figure 14). For this purpose, a reference surface was constructed based on known crane kinematics and results of photogrammetric measurements. All errors were calculated as projection distances to reference surface, however, this approach was used only for overall assessment.

For more accurate analysis it was proposed to calculate four basic types of geometric errors for each set of points. All the recorded data points have been imported to Matlab (The MathWorks, Inc., Natick, USA) and visualized (Figure 15).

As a reference, a set of theoretical motion paths were calculated based on photogrammetric measurements and actual angles of the crane boom read from the attached protractor. The first considered error was the radius error, i.e., the difference between the radius of the circle aligned to the registered points and the radius of the reference circle. Other considered indicators were flatness of the circle aligned to the registered data and parallelism of the aligned circle to the reference

TABLE 2. Geometric Errors for Reconstructed Motion Trajectories of the Crane Tip

No.	Measured boom angle [°]	Reference radius [mm]	Actual radius [mm]	Normalized radius error [%]	Circle flatness error [mm]
1	47.25	3290.3	3 306.8	0.50	12.6
2	52.25	3453.6	3 469.7	0.47	5.3
3	57.00	3582.8	3 598.3	0.43	2.9
4	62.00	3690.4	3 711.7	0.58	2.3
5	67.42	3773.3	3 794.1	0.55	4.2
6	72.00	3815.4	3 836.3	0.55	5.2
7	77.00	3831.8	3 859.5	0.72	8.7
8	82.20	3816.2	3 833.7	0.46	8.6
9	86.80	3773.9	3 803.0	0.77	11.8
10	91.90	3697.4	3 726.1	0.78	10.4
11	96.70	3597.9	3 622.9	0.69	9.9
12	101.90	3457.5	3 476.9	0.56	18.2
13	106.90	3294.9	3 315.3	0.62	26.4
14	111.60	3119.3	3 127.4	0.26	24.6
15	116.20	2921.1	2 931.0	0.34	20.0
16	121.20	2686.0	2 692.6	0.25	20.3

TABLE 3. Geometric Errors for Reconstructed Motion Trajectories of the Crane Tip

No.	Measured boom angle [°]	Maximum deviation to reference plane [mm]	Minimum deviation to reference plane [mm]	Parallelism error [mm]	Circle position error [mm]
1	47.25	4.2	-30.1	34.3	58.6
2	52.25	-3.9	-20.6	24.5	50.6
3	57.00	-10.2	-19.6	29.8	45.3
4	62.00	0.9	-2.1	3.1	29.5
5	67.42	-5.1	-12.4	17.4	20.8
6	72.00	3.9	-2.7	6.6	1.3
7	77.00	17.2	3.6	20.7	22.3
8	82.20	13.6	-7.4	21.0	32.8
9	86.80	28.7	-10.8	39.4	49.3
10	91.90	53.9	-2.5	56.5	66.3
11	96.70	47.8	-16.2	64.0	66.1
12	101.90	75.4	-11.9	87.3	79.6
13	106.90	80.6	-29.3	109.9	81.6
14	111.60	101.5	-36.0	137.5	86.0
15	116.20	114.2	-35.0	149.3	90.8
16	121.20	124.7	-37.1	161.7	83.6

circle. Both the errors have been calculated according to the Geometric Dimensioning and Tolerancing standards (ASME Y14.5-2009). The last considered error was the Cartesian distance between centers of the aligned and the reference circle. All the described errors and their components are presented in Table 2 and Table 3. Measurement no. 6 relates to the initial angle of the boom for which photogrammetric measurement had been carried out.

The normalized percentage radius error was similar in accordance with the change of the boom angle and equals approximately half a percent. The results remained relatively small and acceptable in the considered large workspace. Different relations were observed in the parallelism of reference and actual circles and in the relative distance between the centers of both circles. By changing the boom angle, the smallest error values were observed near the initial angle and increased significantly at the boundaries of measured volume. Differently, the flatness error values of the reconstructed data points were stable for most of the lower boom tip positions and were increasing slightly in the highest positions. The parallelism error in each case has a greater value than the flatness error because geometric tolerance of parallelism contains geometric tolerance of flatness.

It was noticed that as the angle of the boom raises, the recorded points show an increasing tendency to tilt in relation to the reference plane. As a result, in the right part of the workspace, each measured distance tends to be reduced. Similar systematic errors were remarked in the length pattern test, where negative length errors were recorded on the left side of the workspace.

IV. DISCUSSION

The conducted experiments allowed evaluating the accuracy of the non-stationary stereovision system used for three-dimensional tracking of the loading crane. However, the results do not allow to describe the accuracy by one general value. This is due to the impossibility of synchronous recording of reconstructed and reference points one by one, which is caused by the open-loop crane control system. The loading crane and the vision system are two physically separated and independent operating systems, where only the vision system can record actual crane tip positions over time. Therefore, only the theoretical circular motion paths could have been compared with the recorded points. Nevertheless, extended evaluation of the system is possible through proper interpretation of several described errors.

The last two errors from Table 3 seem to have the most impact on the overall measuring error. However, all the observed tendencies, as well as the obtained error distribution map (Figure 19), showed room for possible improvements. Inaccuracies in reconstruction may be caused by incorrectly calibrated rotation axes of each gimbal in reference to the crane origin. The problem may arise from Step 2 of the system calibration, where only a small number of points is used to fit the circle models and determine the rotation axes, which in turn may cause incorrect relation of gimbal and camera coordinate systems.

Similar results were obtained in another research, where the laser system has been used to measure the distance and control position of the crane hook [20]. The tracking error was less than 50 mm in a distance up to 23 m. However, the performance and abilities of the non-stationary stereovision system proposed in this paper are much higher, since the system can observe and track several objects and reconstruct actual crane configuration at the same time. Moreover, in the presented system, no precise camera tracking is needed which simplifies the gimbal control algorithm.

The use of the real crane for the tests was the technical merit of the entire experiment, however, it also caused

some difficulties. Apart from the errors of the vision system itself, the obtained results could also depend on mechanical issues such as various deflection of the crane structure and its mounting base while lifting of the boom or rotating from one side to the other [49]–[53]. Thereby, models of reference circles, which are based only on known kinematics of the crane and single photogrammetric measurement for the initial position of the crane tip, may have been determined inaccurate. Moreover, the gimbals are individually situated on the floor and separated from the crane base. It is worth mentioning that in any circumstances the crane has a much higher static susceptibility under load and higher dynamic susceptibility while accelerating the crane than the observed errors.

V. CONCLUSION AND FUTURE WORK

The non-stationary stereo vision system, composed of two independent gimbals mounted outside the crane structure, enables efficient tracking and reconstruction of the position and configuration of the entire crane. It is possible to track multiple points simultaneously, including the load or obstacles, as long as they are visible in cameras images. Multistage calibration is required after the first assembly of the system. Normal operation requires continuous recalculation of the cameras positions and orientations, based on current encoders positions and well-known gimbals kinematics. The presented system is flexible in terms of calibration and allows adapting to different configurations in the working area.

The first experiment demonstrated typical reconstruction errors of stereo vision systems. It was due to the observation of the crane tip at a high angle at the boundaries of the workspace. There were also systematically observed tendencies to deform the space containing reconstructed points, which suggested incorrectly performed calibration of the gimbals mechanical axes related to camera coordinate systems. The errors could also be affected by the assumed measurement method based on theoretical motion paths of the crane tip, unstable mounting system, and unknown deflection of the entire crane structure during rotation around the Z-axis.

By introducing several improvements and more accurate gimbals calibration, the system would be far enough to measure the crane position in its working space with more than required accuracy for normal handling. A new spatial calibration pattern needs to be designed with calibration points arranged around the camera, and not just in front of it. Therefore, rotation of gimbal axes could be performed in a wider range with the calibration pattern constantly visible in the camera field of view. For a better comparison of real and reconstructed motion paths an additional measuring equipment must be utilized, e.g., laser tracker.

The proposed system has been successfully developed to solve the inverse kinematics problem in real-time, which in the upcoming research may enable to control the crane with a suspended load. Besides, the future design of lightweight gimbals construction will be desirable to improve the overall system abilities and dynamics.

REFERENCES

[1] S. Weyer, M. Schmitt, M. Ohmer, and D. Gorecky, "Towards industry 4.0–standardization as the crucial challenge for highly modular, multi-vendor production systems," *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 579–584, 2015, doi: 10.1016/j.ifacol.2015.06.143.

[2] V. Paelke, "Augmented reality in the smart factory: Supporting workers in an industry 4.0. environment," in *Proc. IEEE Emerg. Technol. Factory Automat.*, Jan. 2014, pp. 1–4, doi: 10.1109/ETFA.2014.7005252.

[3] L. Pérez, Ì. Rodríguez, N. Rodríguez, R. Usamentiaga, and D. García, "Robot guidance using machine vision techniques in industrial environments: A comparative review," *Sensors*, vol. 16, no. 3, p. 335, Mar. 2016, doi: 10.3390/s16030335.

[4] R. Spangenberg, T. Langner, S. Adfeldt, and R. Rojas, "Large scale semi-global matching on the CPU," in *Proc. IEEE Intell. Vehicles Symp. Proc.*, Jun. 2014, pp. 195–201, doi: 10.1109/IVS.2014.6856419.

[5] D. Hernandez-Juarez, A. Chacón, A. Espinosa, D. Vázquez, J. C. Moure, and A. M. López, "Embedded real-time stereo estimation via semi-global matching on the GPU," *Procedia Comput. Sci.*, vol. 80, pp. 143–153, 2016, doi: 10.1016/j.procs.2016.05.305.

[6] V. Kolmogorov and R. Zabih, "Graph cut algorithms for binocular stereo with occlusions," in *Handbook of Mathematical Models in Computer Vision*. New York, NY, USA: Springer, 2006, pp. 423–437, doi: 10.1007/0-387-28831-7_26.

[7] O. Rahnama, T. Cavalleri, S. Golodetz, S. Walker, and P. Torr, "R3SGM: Real-time raster-respecting semi-global matching for power-constrained systems," in *Proc. Int. Conf. Field-Program. Technol. (FPT)*, Dec. 2018, pp. 105–112, doi: 10.1109/FPT.2018.00025.

[8] M. P. Muresan, S. Nedevschi, and R. Danescu, "A multi patch warping approach for improved stereo block matching," in *Proc. 12th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, 2017, pp. 459–466, doi: 10.5220/0006134104590466.

[9] H. Alismail, B. Browning, and S. Lucey, "Photometric bundle adjustment for vision-based SLAM," in *Proc. Asian Conf. Comput. Vis.*, vol. 10114, Nov. 2017, pp. 324–341, doi: 10.1007/978-3-319-54190-7_20.

[10] J. Iglhaut, C. Cabo, S. Puliti, L. Piermattei, J. O'Connor, and J. Rosette, "Structure from motion photogrammetry in forestry: A review," *Current Forestry Rep.*, vol. 5, no. 3, pp. 155–168, Sep. 2019, doi: 10.1007/s40725-019-00094-3.

[11] F. Chiabrando, E. Donadio, and F. Rinaudo, "SfM for orthophoto to generation: A winning approach for cultural heritage knowledge," *ISPRS - Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vols. 5, pp. 91–98, Aug. 2015, doi: 10.5194/isprsarchives-XL-5-W7-91-2015.

[12] G. Caroti, I. Martínez-Espejo Zaragoza, and A. Piemonte, "Accuracy assessment in structure from motion 3D reconstruction from UAV-born images: The influence of the data processing methods," *Proc. Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vols. 4, pp. 103–109, Aug. 2015, doi: 10.5194/isprsarchives-XL-1-W4-103-2015.

[13] F. Zhong, R. Kumar, and C. Quan, "A cost-effective single-shot structured light system for 3d shape measurement," *IEEE Sensors J.*, vol. 19, no. 17, pp. 7335–7346, Sep. 2019, doi: 10.1109/JSEN.2019.2915986.

[14] J. O'Connor, "Impact of image quality on SfM Photogrammetry: Colour, compression and noise," Ph.D. dissertation, Dept. Sci., Eng. Comput., Kingston Univ., London, U.K., 2018.

[15] L. Zhang, K. Xu, S. Yu, R. Fu, and Y. Xu, "An effective approach for active tracking with a PTZ camera," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, Dec. 2010, pp. 1768–1773, doi: 10.1109/ROBIO.2010.5723599.

[16] R. J. Rajesh and P. Kavitha, "Camera gimbal stabilization using conventional PID controller and evolutionary algorithms," in *Proc. Int. Conf. Comput., Commun. Control*, Jan. 2016, pp. 1–6, doi: 10.1109/IC4.2015.7375580.

[17] C. Kanellakis and G. Nikolakopoulos, "Survey on computer vision for UAVs: Current developments and trends," *J. Intell. Robot. Syst.*, vol. 87, no. 1, pp. 141–168, Jul. 2017, doi: 10.1007/s10846-017-0483-z.

[18] T. Luhtmann, S. Robson, S. Kyle, and I. Harley. (2011). *Close Range Photogrammetry: Principles, Techniques and Applications*. Accessed: Jul. 22, 2020. [Online]. Available: https://www.researchgate.net/publication/237045019_Close_Range_Photogrammetry_Principles_Techniques_and_Applications

[19] K. Middlicki and M. Saków, "LiDAR based system for tracking loader crane operator," in *Advances in Manufacturing* (Lecture Notes in Mechanical Engineering). London, U.K.: Pleiades, 2019, pp. 406–421.

[20] K. Miadlicki, M. Pajor, and M. Sakow, "Ground plane estimation from sparse LIDAR data for loader crane sensor fusion system," in *Proc.*

22nd Int. Conf. Methods Models Autom. Robot. (MMAR), Aug. 2017, pp. 717–722, doi: 10.1109/MMAR.2017.8046916.

[21] K. Miadlicki, M. Pajor, and M. Sakow, "Real-time ground filtration method for a loader crane environment monitoring system using sparse LIDAR data," in *Proc. IEEE Int. Conf. Innov. Intell. Syst. Appl. (INISTA)*, Jul. 2017, pp. 207–212, doi: 10.1109/INISTA.2017.8001158.

[22] K. Miadlicki, M. Pajor, and M. Sakow, "Loader crane working area monitoring system based on lidar scanner," in *Advances in Manufacturing* (Lecture Notes in Mechanical Engineering), no. 201519. 2018, pp. 465–474, doi: 10.1007/978-3-319-68619-6_45.

[23] O. Naroditsky, A. Patterson, and K. Daniilidis, "Automatic alignment of a camera with a line scan LIDAR system," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 3429–3434, doi: 10.1109/ICRA.2011.5980513.

[24] B. Langmann, K. Hartmann, and O. Loffeld, "Depth camera technology comparison and performance evaluation," in *Proc. IEEE 1st ICPRAM*, Portugal, Lisbon, vol. 2, Feb. 2012, pp. 438–444.

[25] M. Sakow, A. Parus, M. Pajor, and K. Miadlicki, "Unilateral hydraulic telemanipulation system for operation in machining work area," in *Advances in Manufacturing* (Lecture Notes in Mechanical Engineering), no. 201519. 2018, pp. 415–425, doi: 10.1007/978-3-319-68619-6_40.

[26] M. Sakow, A. Parus, M. Pajor, and K. Miadlicki, "Nonlinear inverse modeling with signal prediction in bilateral teleoperation with force-feedback," in *Proc. 22nd Int. Conf. Methods Models Autom. Robot. (MMAR)*, Aug. 2017, pp. 141–146, doi: 10.1109/MMAR.2017.8046813.

[27] K. Stateczny, P. Herbin, and M. Pajor, "Six-axis control joystick based on tensometric beam," *Adv. Manuf. Sci. Technol.*, vol. 40, pp. 33–41, 2016, doi: 10.2478/amst-2016-0020.

[28] P. Herbin and M. Pajor, "ExoArm 7-DOF (interactive 7-DOF motion controller of the operator arm) master device for control of loading crane," in *Advances in Manufacturing* (Lecture Notes in Mechanical Engineering), no. 201519. 2018, pp. 439–449, doi: 10.1007/978-3-319-68619-6_42.

[29] K. C. C. Peng, W. Singhose, and P. Bhaumik, "Using machine vision and hand-motion control to improve crane operator performance," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 42, no. 6, pp. 1496–1503, Nov. 2012, doi: 10.1109/TSMCA.2012.2199301.

[30] M. Majewski, W. Kacalak, Z. Budniak, and M. Pajor, "Interactive control systems for mobile cranes," in *Proc. Adv. Intell. Syst. Comput.*, vol. 661, 2018, pp. 10–19, doi: 10.1007/978-3-319-67618-0_2.

[31] D. Miyamoto, S. Nara, and S. Takahashi, "Visual feedback control with laser for the position detection of crane hook," in *Proc. SICE-ICASE Int. Joint Conf.*, vol. 2006, pp. 2079–2083, doi: 10.1109/SICE.2006.315555.

[32] S. Nara, D. Miyamoto, S. Takahashi, and S. N. I. Kaneko, "Visual feedback tracking of crane hook," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, May 2006, pp. 2736–2742, doi: 10.1109/IROS.2006.281999.

[33] M. Majewski and W. Kacalak, "Human-machine speech-based interfaces with augmented reality and interactive systems for controlling mobile cranes," in *Proc. Int. Conf. Interact. Collaborative Robot.*, vol. 9812, 2016, pp. 89–98, doi: 10.1007/978-3-319-43955-6_12.

[34] P. Hyla, "The crane control systems: A survey," in *Proc. 17th Int. Conf. Methods Models Autom. Robot.*, 2012, pp. 505–509, doi: 10.1109/MMAR. 2012.6347867.

[35] Y. Yoshida, "Gaze-controlled stereo vision to measure position and track a moving object: Machine vision for crane control," in *Smart Sensors, Measurement and Instrumentation*, vol. 8. Springer, 2014, pp. 75–93, doi: 10.1007/978-3-319-02315-1_4.

[36] A. M. Aurand, J. S. Dufour, and W. S. Marras, "Accuracy map of an optical motion capture system with 42 or 21 cameras in a large measurement volume," *J. Biomechan.*, vol. 58, pp. 237–240, Jun. 2017, doi: 10.1016/j.jbiomech.2017.05.006.

[37] P. Merriaux, Y. Dupuis, R. Boutteau, P. Vasseur, and X. Savatier, "A study of vicon system positioning performance," *Sensors*, vol. 17, no. 7, p. 1591, Jul. 2017, doi: 10.3390/s17071591.

[38] F. Chen, X. Chen, X. Xie, X. Feng, and L. Yang, "Full-field 3D measurement using multi-camera digital image correlation system," *Opt. Lasers Eng.*, vol. 51, no. 9, pp. 1044–1052, Sep. 2013, doi: 10.1016/j.optlaseng.2013.03.001.

[39] K. Pra?nowski, A. Bieniek, J. Mamala, and K. Hennek, "Using the Pontos optical system for deformation's measurement of the vehicle tyre at dynamic conditions," *AIP Conf. Proc.*, vol. 2029, Oct. 2018, Art. no. 020060, doi: 10.1063/1.5066522.

[40] A. Klimchik and A. Pashkevich, "Robotic manipulators with double encoders: Accuracy improvement based on advanced stiffness modeling and intelligent control," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 740–745, 2018, doi: 10.1016/j.ifacol.2018.08.407.

[41] M. Pajor, M. Grudziński, and L. Marchewka, "Stereovision system for motion tracking and position error compensation of loading crane," *AIP Conf. Proc.*, vol. 2029, Dec. 2018, Art. no. 020050, doi: 10.1063/1.5066512.

[42] T. Luhmann, C. Fraser, and H. G. Maas, "Sensor modelling and camera calibration for close-range photogrammetry," *ISPRS J. Photogramm. Remote Sens.*, vol. 115. Amsterdam, The Netherlands: Elsevier, May 2016, pp. 37–46, doi: 10.1016/j.isprsjprs.2015.10.006.

[43] J.-Y. Bouguet. (2003). *Camera Calibration Toolbox for MATLAB*. Accessed: Jul. 22, 2020. [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc/index.html

[44] F. Remondino. (2006). *Detectors and Descriptors for Photogrammetric Applications*. [Online]. Available: http://dx.doi.org/10.1.1.71.3114

[45] S. Ji, X. Fan, D. W. Roberts, A. Hartov, and K. D. Paulsen, "Flow-based correspondence matching in stereovision," in *Proc. Int. Workshop Mach. Learn. Med. Imag.*, vol. 8184, vol. 2013, pp. 106–113, doi: 10.1007/978-3-319-02267-3_14.

[46] J. Wang, E. Kobayashi, and I. Sakuma, "Coarse-to-fine dot array marker detection with accurate edge localization for stereo visual tracking," *Biomed. Signal Process. Control*, vol. 15, pp. 49–59, Jan. 2015, doi: 10.1016/j.bspc.2014.09.008.

[47] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proc. 7th IEEE Int. Conf. Comput. Vis.*, Dec. 1999, pp. 666–673, doi: 10.1109/iccv.1999.791289.

[48] M. Grudziski and A. Marchewka, "A stereovision system for three-dimensional measurements of machines," *Sci. Journals Marit. Univ. Szczecin*, vol. 58, no. 130, pp. 16–23, 2019, doi: 10.17402/332.

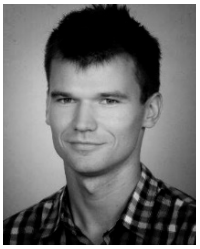
[49] W. Kacalak, Z. Budniak, and M. Majewski, "Stability assessment as a criterion of stabilization of the movement trajectory of mobile crane working elements," *Int. J. Appl. Mech. Eng.*, vol. 23, no. 1, pp. 65–77, Feb. 2018, doi: 10.1515/ijame-2018-0004.

[50] P. Dunaj, B. Niesterowicz, and B. Szymczak, "Loader crane modal analysis using simplified hydraulic actuator model," in *Proc. Int. Sci.-Tech. Conf.* London, U.K.: Pleiades, 2019, pp. 70–80.

[51] S. Duda, K. Kawlewski, and G. Gembalczyk. (2015). *Concept of the System for Control Over Keeping Up the Movement of a Crane*. Accessed: Jul. 22, 2020. [Online]. Available: https://www.researchgate.net/publication/273129408_Concept_of_the_System_for_Control_over_Keeping_up_the_Movement_of_a_Crane

[52] W. Kacalak, Z. Budniak, and M. Majewski, "Computer aided analysis of the mobile crane handling system using computational intelligence methods," *Adv. Intell. Syst. Comput.*, vol. 662, pp. 250–261, 2018, doi: 10.1007/978-3-319-67621-0_23.

[53] S. Såawski, M. Szymiczek, J. Kaczmarczyk, J. Domin, and S. Duda, "Experimental and numerical investigation of striker shape influence on the destruction image in multilayered composite after low velocity impact," *Appl. Sci.*, vol. 10, no. 1, p. 288, Dec. 2019, doi: 10.3390/app10010288.



MAREK GRUDZIŃSKI was born in Szczecin, Poland, in 1985. At the age of 19, he began studying. He received the M.A. degree in automation and robotic systems from the West Pomeranian University of Technology, Szczecin, and the Ph.D. degree from the Department of Mechanical Engineering and Mechatronics, West Pomeranian University of Technology, in 2015. His interests were focused mainly on automation of manufacturing processes and optical metrology, therefore in 2009, he began his doctoral studies. Since 2015, he has been employed at the same department as a Professor's Assistant and Lecturer. At the same time, he was involved in four scientific research in cooperation with local industry. His research interests include optical measurement techniques, stereovision based systems, ophthalmic devices, images processing, augmented reality, as well as mechatronic solutions such as autonomous flying robots and CNC control systems. He is the coauthor of more than 20 articles and conference speech and hold three patents.

ŁUKASZ MARCHEWKA was born in Szczecin, Poland, in 1995. At the age of 19, he had graduated high school. He received the degree (Hons.) in mechatronics engineering from the Faculty of Mechanical Engineering and Mechatronics, West Pomeranian University of Technology, Szczecin, and the master's degree in mechanical engineering from the West Pomeranian University of Technology, in 2019. He is currently pursuing the Ph.D. degree. For three years, he has projected and developed vision systems. Moreover, for three years he was working in a research project in cooperation with Cargotec company, funded by the National Center for Research and Development. The result of participation in this project was two published scientific articles and also admittance to Doctoral School.

MIROSŁAW PAJOR graduated from the Faculty of Mechanical Engineering, Szczecin University of Technology, in 1991. He received the Ph.D. degree, in 1997, and the postdoctoral degree in the discipline of machine construction and exploitation. in 2006. He is currently the Dean of the Faculty of Mechanical Engineering and Mechatronics for the second term (2020–2024). He specializes in modeling and research of machining processes, dynamics of machine tools and robots, designing of machine diagnostic systems and mechatronics. He published 126 works. In the last six years he managed five research projects carried out in cooperation with industrial partners. He promoted four doctors and currently cooperate with two doctoral students.

RYSZARD ZIĘTEK was born in Szczecin, Poland, in 1995. At the age of 19, he began studying. He received the degree in mechatronics engineering from the Faculty of Mechanical Engineering and Mechatronics, West Pomeranian University of Technology, Szczecin, and the master's degree in mechanical engineering from the West Pomeranian University of Technology, in 2020. For three years, he has developed systems based on field-programable technology. He is currently a candidate for the Doctoral School.