

Design of a Motion Monitoring System for Unmanned Offshore Topside Installations based on real-time Visual Object Tracking using Drones and Fixed Cameras on an SSCV

V.E. (Vincent) Mullenders¹

Abstract— This paper presents a novel approach for estimating the relative motion between a moving topside and a jacket. The method is based on an algorithm including a multi-camera vision system. The vision system makes use of the open-source computer vision library OpenCV and planar Aruco markers placed on the topside and jacket. Aruco was only recently introduced which makes this solution unique in the offshore sector. The Motion Tracking Algorithm is designed in such a way that it is not required to define the position of the cameras. The imaging devices includes fixed cameras on a heavy lift vessel or cameras attached under a drone. Experimental results in a realistic virtual environment demonstrate the accuracy and efficiency of the proposed method. Both the mean absolute error and standard deviation error was found to be 48 mm, or better when different camera configurations are used.

I. INTRODUCTION

As well known in the offshore oil and gas industry, an offshore platform consists of two parts structurally the upper part or the topside and the lower part or the jacket. A topside generally is a steel structure consisting of more than one deck holding various kinds of facilities for exploration or production. In the offshore oil and gas industry, the completion of a platform construction must go through the step of the integration of the topside with the corresponding jacket. The topside can be installed by a single lift using a Semi Submersible Crane Vessel (SSCV).

Setting down a topside on a jacket is considered as a critical offshore installation activity. A high level of concentration and perfect communication between parties involved is necessary during this operation. Prior to positioning the topside, the rigger foreman and assistant superintendent take place on the jacket. With the role to communicate topside positioning information to the superintendent - who is in charge of the operation - on the crane vessel. Despite the fact that incidents are extremely rare, it is an unwanted situation to have people on the jacket. Besides the dangers of standing under a suspended load [4] of up to 10,000mt the need for crew transfer between the vessel and jacket is also eliminated.

Current efforts from stakeholders in the offshore industry to enable unmanned topside installations are based on Laser Measurements and Vision Based Tracking of natural features (real world objects). The laser based system uses robotic 'Total Stations' to simultaneously track 360 degree prisms

on the topside and jacket. A computing device is then able to calculate the relative position. The Vision Based system uses images from two cameras to match the filmed structures with a predefined 3-D CAD model of the topside and jacket in real-time. Such a system is vital to the success of a multi-million dollar project. Accuracy and precision is therefore a very important pillar.

A. Concept 1: Robotic Total Stations

Concept 1 is a solution that uses six robotic total stations to track six prisms and is shown in Figure 1.



Fig. 1: Robotic Total Stations

These prisms are located on the topside and jacket. Three total stations track three prisms on the jacket and three total stations track three prisms on the topside. Information from the six total stations is transferred to a computing device to calculate the relative position between the topside and the jacket. This information can be displayed in real-time. The total stations are located on the stern of the SSCV and their position with respect to the vessel reference frame must be precisely calibrated. This is also necessary for the six prisms. Their 3-D location with respect to the topside or jacket reference frame must be obtained during a dimensional control survey. For the topside, this can be performed onshore at the yard. The prisms on the jacket must be installed offshore prior to the topside installation. The positioning method of total stations is based on a three-dimensional coordinator measuring technology. The accuracy of this system is within 0.057 m standard deviation for the combined 3-D position. And within 0.066 degrees standard deviation for combined pitch and roll. The accuracy of this system

*This work was supported by Heerema Marine Contractors

¹V.E. Mullenders is with Faculty of Mechanical, Maritime and Materials Engineering, Technical University of Delft, 2628 CN Delft, The Netherlands

is highly dependent on the precision of the dimensional control survey. This system has been used several times as the primary system for unmanned topside installations. The system works well but requires a lot of preparation to set up. In addition, one does not yet dare to trust this system blindly. Some feedback from a superintendent during an unmanned topside installation was:

- A camera to confirm what you see would be beneficial
- It is still difficult to keep the complete overview when looking at a screen.
- Some kind of visual information would be nice.
- It feels much better to rely on a system which has been confirmed by the crew below the load or my own visual interpretation of reality.

B. Concept 2: Augmented Reality System

Concept 2 uses cameras to detect, recognize and track objects by means of augmented reality (AR). AR can match and track real world objects and project 3-D predefined CAD drawings on top of it. This is shown in Figure 2.

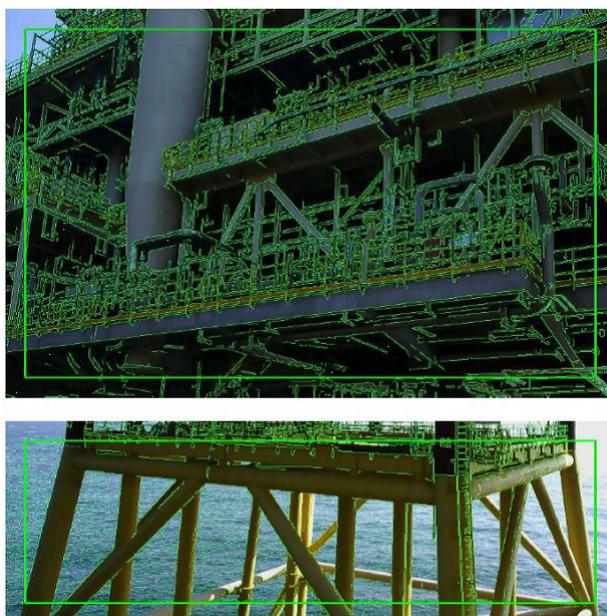


Fig. 2: Augmented Reality System

A camera is used to film the topside and jacket, the green lines indicate a successful recognition and tracking. For every installation a 3-D CAD model must be prepared in order to facilitate object tracking. The two cameras are connected to a computing device to determine relative position. Despite many efforts this system never worked in an offshore environment. Shadow lines on the structures confused the system which resulted in inaccurate tracking. No statement can be made about the expected accuracy and precision.

C. Comparison between concept 1 and 2

In the process of engineering a new system it is useful to compare the predecessor systems. The predecessor systems can serve as a point of departure. It is the single richest

source of information on the requirements for a new system. Both concepts were developed to support unmanned installation of topsides. Concept 1, with the robotic total stations, has been used several times as primary system but requires a lot of preparation and is quite expensive, ranging from 75K - 100K USD per installation. Concept 2, the augmented reality system, had a lot of potential but failed to work properly in the offshore environment. There is still a need for a simpler and cost-efficient system that can be operated by HMC itself. The advantage of concept 2 was that a camera was used as the primary sensor. This allows the superintendent to clearly see what happens and interpret the acquired data clearly. For concept 1, the superintendent did not dare to rely entirely on the system and a second system with cameras was a desired backup. A limitation of both systems is that there is not always enough visibility on the jacket from the SSCV. This is the case, for example, if the jacket does not project far above the water level. A system with flexible cameras would therefore be favorable.

D. System requirements

During the development of a new concept, for the monitoring of offshore topsides during installation, the following requirements were taken into account:

- The proposed solution should accommodate any type of topside as well as any installation method (single or double crane). The system shall function irrespective of vessel type, sea fastening or vessel positioning system.
- Concentricity must be measured with a precision of below 150 mm.
- The superintendent (located on the stern of the SSCV during the installation) is the decision maker during the installation process. Any information on the positioning of the topside should be simply and clearly transmitted to him.
- Robustness is an important pillar as the system shall perform when needed. Should the primary system fail, a back-up system/solution needs to be implemented.
- No further risk shall be added to the installation process.
- Should weather conditions become too dangerous to carry on, the installation can be postponed. This can significantly lengthen the whole installation from lift-off to set-down. Any system should work continuously for up to six hours.
- Any solution should not only work in perfect conditions, but also at the limits to what is still considered a safe situation (where significant movement and impact can be expected). Wind speeds up to 30 knots (15.34 m/s), sea water exposure, humidity and temperatures of between -10°C to 40°C should not effect the system and its output.

E. New concept: Visual object tracking with fixed cameras and drones

Following these approaches, the work presented herein aims to show that an image-based method with multiple cameras can monitor the relative topside movements during a

virtually simulated topside installation. In this paper, a novel motion tracking algorithm is presented based on drones, fixed cameras and visual object tracking (Fig. 3). Drones



Fig. 3: Drones as an installation aid

are already starting to change how businesses operate and this is happening today. Companies across industries are using them for inspection, monitoring, repair work and onsite security. They are also being used for real-time data collection. Drones are able to take any position with respect to the topside or jacket and can mimic the view from people on the jacket. They are therefore not limited by the view from the SSCV. A limitation in the use of drones is the requirement of a certified operator and the limited power supply. To eliminate the requirement of a certified operator, autonomous drones could be used. Several studies show that there are a lot of opportunities in enabling autonomous drone operations. [9] It will be a matter of time before the first autonomous drones will take off from vessels to carry out specific missions. The power supply of drones is also likely to increase due to a technology push from the automotive and electronic consumer goods industry. [10] The developed algorithm is able to localise a pair of Aruco markers in an image captured by the vision system. Aruco was only recently introduced which makes this solution unique in the offshore sector. If a marker pair is recognized successfully, relative distance calculations can be made. By conveniently placing these markers on the topside stabbing cones and jacket legs, the topside relative motions can be estimated. A minimum of two locations need to be monitored in order to perform a successful estimation. The proposed method is based on the concept of planar homography between the camera CCD plane and a pair of ArUco markers [2] fixed to the body under inspection. Briefly, the planar homography relates the transformation between two planes. Accurate 3-D motion tracking through homography is a key problem in Computer Vision in general environments where planar surfaces (e.g., braces, beams of offshore structures) allow additional constraints to the camera pose [6]. The problem is overcome here, since a planar target of known geometry is secured firmly on the topside and jacket. This allows one to define the 3-D relative position between these two structures. In this paper, the image processing framework is presented in the context of scaled models of a topside, a jacket, and a Semi Submersible Crane Vessel (SSCV) in a virtual environment. Performance of the constructed motion tracking algorithm is presented, with an analysis of expected accuracies and resolutions. Finally, examples of captured motion and concluding remarks are presented.

II. METHOD

The proposed optical method is based on a multiple camera view of known flat targets (for example an ArUco marker) fixed on the bodies (jacket and topside) of which 3-D motion is required (Fig. 4).

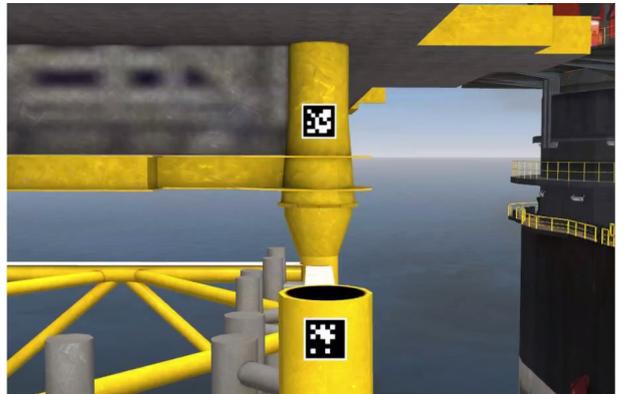


Fig. 4: Jacket and topside with fixed ArUco marker, dimensions 1x1 m.

Targets on the same body must be mounted as good as perpendicular against each other. A marker pair is defined as a set of two markers in the same reference plane. The relative horizontal and vertical motion is calculated using a reference system defined on such a panel. In the proposed method, time series of movements of the bodies are measured using images from the multiple camera setup, that is a sequence of images at a given frame rate. Images must show a flat target of given geometry whose movements in time is followed. The basic idea underlying the method is that for every image of the sequence, it is possible to define the rigid motion $[\mathbf{R} \ \mathbf{t}] \mid \mathbf{R}, \mathbf{t} \in \mathbb{R}^3$ between the camera CCD plane and the pattern plane by exploiting the concept of homography.

In order to estimate the 3-D rigid motion, the camera-lens must be calibrated to calculate the so-called intrinsic camera parameters:

- Lens focal length: f
- Principal points coordinate: c_c
- Image distortion coefficients: k_r (radial)

A. Camera calibration

Camera calibration is the process of obtaining the fundamental parameters of a camera [8]. Calibration only needs to be performed once per camera. These parameters describe the relation between 2D image pixels $\mathbf{u} \in \mathbb{R}^2$ and the real world coordinates $\mathbf{p} \in \mathbb{R}^3$. This relationship is modeled using the pinhole camera model:

$$\begin{bmatrix} x_i \\ y_i \\ w \end{bmatrix} = \underbrace{\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}}_K \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (1)$$

K is known as the camera calibration matrix which can be found from taking several pictures of a checkerboard pattern

and process them using the OpenCV [3] calibration function. In addition to calculate the camera calibration matrix K , it is also necessary to remove radial image distortion. The mathematical relationship between the corrected image pixels (x_i, y_i) and the radial distorted pixels (x_d, y_d) is:

$$x_i = x_d \left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6 \right) \quad (2)$$

$$y_i = y_d \left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6 \right) \quad (3)$$

$$r = \sqrt{x_d^2 + y_d^2} \quad (4)$$

where k_1, k_2, k_3 are the radial distortion coefficients which are also found using the OpenCV calibration function.

B. Three-dimensional transforms and camera model

Consider a three-dimensional point $\mathbf{p}_a = (\mathbf{x}, \mathbf{y}, \mathbf{z})$ in an arbitrary reference frame a . In order to express such point into another reference system b it must undergo a rotation followed by a translation [7]. Let us denote by

$$\zeta = (\mathbf{r}, \mathbf{t}) \mid \mathbf{r}, \mathbf{t} \in \mathbb{R}^3, \quad (5)$$

the three rotational and translational components $\mathbf{r} = (r_x, r_y, r_z)$ and $\mathbf{t} = (t_x, t_y, t_z)$. Using Rodrigues' rotation formula, the rotation matrix \mathbf{R} can be obtained from \mathbf{r} as:

$$\mathbf{R} = \mathbf{I}_{3 \times 3} + \bar{\mathbf{r}} \sin \theta + \bar{\mathbf{r}}^2 (1 - \cos \theta) \quad (6)$$

where $\mathbf{I}_{3 \times 3}$ is the identity matrix and $\bar{\mathbf{r}}$ denotes the antisymmetric matrix

$$\bar{\mathbf{r}} = \begin{bmatrix} 0 & -r_x/\theta & r_y/\theta \\ r_z/\theta & 0 & -r_x/\theta \\ -r_y/\theta & r_x/\theta & 0 \end{bmatrix} \quad (7)$$

such that $\theta = \|\mathbf{r}\|_2$.

Then, in combination with \mathbf{t} , the 4 x 4 matrix

$$\gamma = \Gamma(\zeta) = \begin{bmatrix} \mathbf{R} & \mathbf{t}^T \\ 0 & 1 \end{bmatrix} \quad (8)$$

can be used to transform the point from a to b as:

$$\begin{bmatrix} \mathbf{p}_b^T \\ 1 \end{bmatrix} = \gamma \begin{bmatrix} \mathbf{p}_a^T \\ 1 \end{bmatrix} \quad (9)$$

This notation can be simplified as:

$$\mathbf{p}_b = \gamma \cdot \mathbf{p}_a \quad (10)$$

A point p projects in the camera plane into a pixel $\mathbf{u} \in \mathbb{R}^2$. Assuming that the camera parameters are known (as is in the case of calibrated cameras), the projection can be obtained as a function:

$$\mathbf{u} = \Psi(\delta, \gamma, \mathbf{p}) \quad (11)$$

where

$$\delta = (f_x, f_y, c_x, c_y, k_1, \dots, k_n)$$

refers to the camera intrinsic parameters, comprised by the focal distances (f_x, f_y) , optical center (c_x, c_y) and distortion parameters (k_1, \dots, k_n) . The parameter γ represents camera pose from a single camera frame. It transforms a point from an arbitrary reference system to the camera one.

C. Relative position estimation

If a camera detects a marker pair it is able to determine the 2-D relative horizontal and vertical position. With a second camera perpendicular to the first camera filming a second marker pair on the same target it is possible to reconstruct the 3-D relative position. This is shown for one camera in Figure 5. Each marker returns its unique ID to identify the marker and a joint rotation-translation matrix. This matrix is used to describe the camera motion around a static scene, or vice versa, rigid motion of an object in front of a fixed object. The constructed algorithm is able to find ${}^C\mathbf{M}_S$ and ${}^C\mathbf{M}_W$ as described in Figure 5.

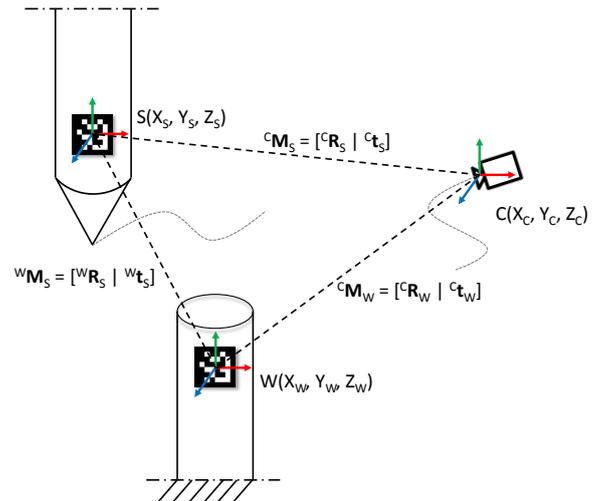


Fig. 5: Kinematic model

These two joint rotation-translation matrices can be used to derive the joint rotation-translation matrix ${}^W\mathbf{M}_S$ between the two bodies, in this case the jacket and topside:

$${}^W\mathbf{M}_S = [{}^W\mathbf{R}_S \mid {}^W\mathbf{t}_S] \quad (12)$$

The rotation matrix ${}^W\mathbf{R}_S$ can be derived from ${}^C\mathbf{R}_W$ and ${}^C\mathbf{R}_S$ as:

$${}^W\mathbf{R}_S = ({}^C\mathbf{R}_W^{-1} \cdot {}^C\mathbf{R}_S)^{-1} \quad (13)$$

The translation vector ${}^W\mathbf{t}_S$ can be derived from ${}^C\mathbf{R}_S$, ${}^C\mathbf{t}_W$, and ${}^C\mathbf{t}_S$:

$${}^W\mathbf{t}_S = {}^C\mathbf{R}_W^{-1} \cdot {}^C\mathbf{t}_S + {}^W\mathbf{R}_C \cdot -{}^C\mathbf{t}_W \quad (14)$$

Since the inverse of a rotation matrix is its transpose and the inverse of a translation vector can be obtained by reversing its direction, equations (13) and (14) can be simplified:

$${}^W\mathbf{R}_S = ({}^C\mathbf{R}_W^{-1} \cdot {}^C\mathbf{R}_S)^{-1} = ({}^W\mathbf{R}_C \cdot {}^C\mathbf{R}_S)^{-1} \quad (15)$$

$${}^W\mathbf{t}_S = {}^C\mathbf{R}_W^{-1} \cdot {}^C\mathbf{t}_S + {}^W\mathbf{R}_C \cdot -{}^C\mathbf{t}_W = {}^W\mathbf{R}_C \cdot {}^C\mathbf{t}_S + {}^W\mathbf{t}_C \quad (16)$$

The final obtained vector ${}^W\mathbf{t}_S$ includes information about the vertical offset $(\Delta x, \Delta y, \Delta z)$ between a pair of markers as can be seen in Figure 6.

An accuracy analysis performed by the authors of [1] showed that there is an error dependence on the distance from the camera to the markers. The further a marker is

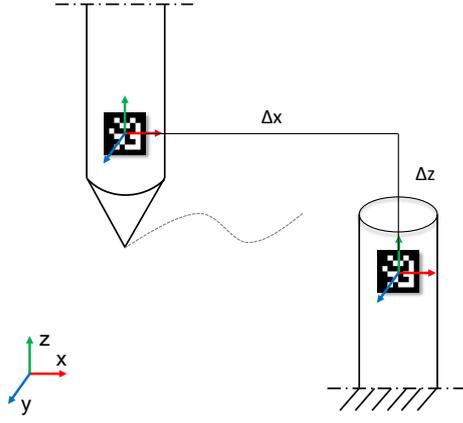


Fig. 6: Kinematic model, Δx -, Δz -offset

from the camera (or the lower the ratio between the marker pixel area and the frame pixel area), the bigger the error in 3-D estimations. This error is mainly present on the axis parallel to the direction of the camera (the depth). On the remaining two axes perpendicular to the camera, the error was negligible. This result is expected since the measurement accuracy in computer vision systems depends on the image pixel density. Therefore, depth measurements contains more noise since small changes in depth will cause the marker edges to move across more pixels in the image if compared to vertical and horizontal movements. For the best possible result, it is therefore best to use two cameras that are perpendicular to each other.

D. Real-time motion tracking

The developed motion tracking algorithm is able to measure the relative 3-D motion between two bodies. In case of an Offshore Topside Installation it is the most convenient way to track the stabbing cones. Camera imagery can come from fixed cameras on the SSCV or from drones [5]. It is recommended to use a frame rate of at least 10 FPS for a continuous signal and a resolution of at least 640x480 pixels. [2] The frames from two cameras are simultaneously analyzed and combined to calculate the 3-D motions. Each frame is converted into a grayscale. Grayscale images require less computing power than colored RGB-images. Detected markers are compared with the marker ID's in a database. If a set of markers is found their pose and relative position is calculated. This results in a Δx - and Δz -offset from camera 1 and an Δy - and Δz -offset from camera 2. Δz is returned twice. A running average of these offsets with a length of 10 samples is calculated to allow for some real-time noise filtering. It is calculated as:

$$\Delta \bar{x} = \frac{\Delta x_M + \Delta x_{M-1} + \dots + \Delta x_{M-(n-1)}}{n} = \frac{1}{n} \sum_{n-1}^{i=0} \Delta x_{M-1} \quad (17)$$

$$\Delta \bar{y} = \frac{\Delta y_M + \Delta y_{M-1} + \dots + \Delta y_{M-(n-1)}}{n} = \frac{1}{n} \sum_{n-1}^{i=0} \Delta y_{M-1} \quad (18)$$

$$\Delta \bar{z}_1 = \frac{\Delta z_{1,M} + \Delta z_{1,M-1} + \dots + \Delta z_{1,M-(n-1)}}{n} = \frac{1}{n} \sum_{n-1}^{i=0} \Delta z_{1,M-1} \quad (19)$$

$$\Delta \bar{z}_2 = \frac{\Delta z_{2,M} + \Delta z_{2,M-1} + \dots + \Delta z_{2,M-(n-1)}}{n} = \frac{1}{n} \sum_{n-1}^{i=0} \Delta z_{2,M-1} \quad (20)$$

The two $\Delta \bar{z}$ -offsets obtained are averaged:

$$\Delta \bar{z} = \frac{\Delta \bar{z}_1 + \Delta \bar{z}_2}{2} \quad (21)$$

The running average will result in a small time delay. If the images are acquired at 30 FPS then the running average will result in a delay of 1/3 second. The running average ensures that the data is displayed more smoothly and fluctuations are filtered out and highlight longer term trends. Mathematically, a moving average is a type of convolution and so it can be viewed as an example of a low-pass filter. If a smaller time-delay is desirable then the length of the samples can be decreased. This will in return lead to a signal with more noise.

III. EXPERIMENTS

In order to evaluate the accuracy of the Motion Tracking Algorithm, two experiments were carried out. The experiments were carried out in a virtual simulation environment. The simulation environment is shown in Figure 7.



Fig. 7: Simulation setup

A topside and a jacket were equipped with unique markers with dimensions of 1x1 m to enable tracking (see Fig. 4). In the first experiment, the camera viewpoint of the simulator is calibrated. In the second experiment, the accuracy and precision of the Motion Tracking Algorithm is measured. During all experiments the axis definition as described in Figure 8 was used.

A. Experiment 1: Calibration

Camera calibration is a crucial part to enable real-time motion tracking as explained in section II-A. Because a virtual environment is used it is not possible to use physical cameras. Therefore the calibration will be used to determine how the simulation software renders objects in a virtual environment and measures its distortion. For a successful calibration at least 20 images are required which include a chessboard pattern. A set of 30 images (or screen shots) were taken

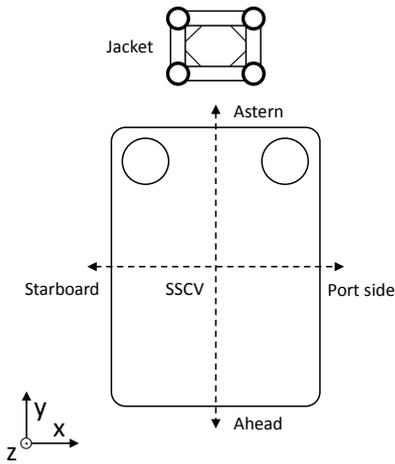


Fig. 8: Axis definition used in the experiments

from different angles. The output of the calibration are the camera- and distortion coefficients collected in the set δ as described in section II-B.

B. Experiment 2: Algorithm accuracy and precision

In this experiment the accuracy and precision of the Algorithm is tested. Accuracy is a very important condition that must be met before the introduction of a new system. Two configurations will be tested. Configuration 1 (Fig. 9) uses four fixed cameras on the SSCV and one drone. Configuration 2 (Fig. 10) uses three drones equipped with a camera and one fixed camera on the stern of the SSCV. These configurations have been chosen to show the versatility of the system. Multiple viewpoints are created in the virtual environment to simulate the three drone and the fixed camera. These viewpoints are the input for the Motion Tracking Algorithm to determine the relative position of the topside with respect to the jacket in real-time. This data is saved to a *.CSV-file and can be compared to the simulation log created during the experiment.

1) *Configuration 1*: In this configuration four cameras are positioned on the SSCV. In this case, the stabbing cones that are closest to the SSCV are monitored (A and B in Fig. 9). If these two are well aligned, the rear cones must also be well aligned. A drone can be used to confirm this. This configuration is shown in Fig. 9. It is also clear that the markers and the cameras are perpendicular to each other, but together make an angle of 45 degrees with respect to the SSCV. In this way, a measurement can be performed from two sides in order to form a three-dimensional image. The topside was moved following different procedures as described in table I. During procedures 1 and 2 the topside was moved by the SSCV. During procedures 3 and 4 the topside was moved due to vessel motions caused by waves defined as a Jonswap spectrum. Also fog and rain was added in the simulation environment to add noise to the camera images.

2) *Configuration 2*: In configuration 2 markers are placed at the primary stabbing cones (A and D in Fig. 10). This

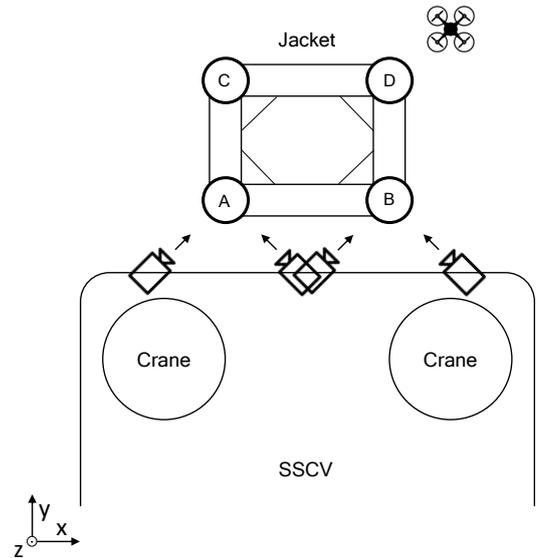


Fig. 9: Configuration 1 - Camera and marker configurations.

TABLE I: Configuration 1: different moves

Procedure	Moves (x, y)	Description
(1)	(0, -20) (0, 0)	Start position 20m ahead, Move 20m astern to final position.
(2)	(0, 0) (10, 0)	Start position above topside, move 10m to port side.
(3)	(0, 0)	Test performance for: Jonswap: 1m, 8s, heading 135, spreading: 4 Fog: 50% Rain: 30%
(4)	(0, 0)	Test performance for: Jonswap: 1m, 8s, heading 135, spreading: 4 Fog: 50% Rain: 30%

is similar to the current procedure whereby the assistant superintendent and rigger foreman are positioned next to primary and secondary stabbing cones. Markers on stabbing cone A are filmed by a drone and a fixed camera at the stern of the SSCV. The markers on stabbing cone D are filmed by two drones. This configuration is schematically shown in Figure 10.

Each marker pair has one marker on the jacket (bottom) and one marker on the topside (top). The topside was moved in three different procedures by the SSCV. These procedures are given in table II.

IV. RESULTS

A. Experiment 1: Calibration

Normally camera calibration is performed by taking a set of pictures of a calibration board. In this case that was not possible as no physical cameras were used. During the calibration experiment a calibration board with a chessboard pattern was inserted in the virtual environment. The camera viewpoint was then adjusted to take several screen shots from

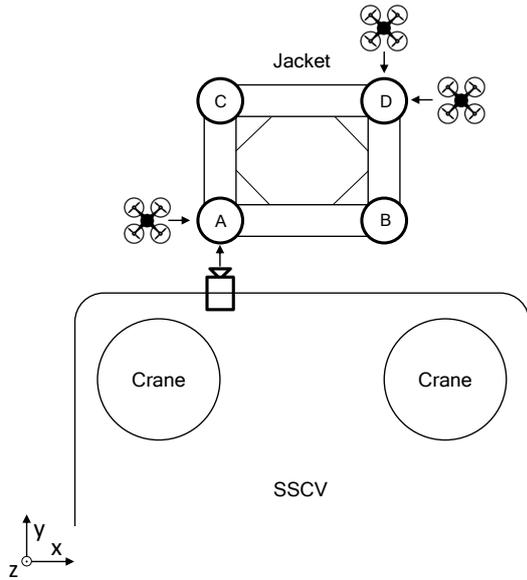


Fig. 10: Configuration 2 - Camera and marker configurations.

TABLE II: Configuration 2: different moves

Procedure	Moves (x, y)	Description
(5)	(0, -5)	Start position 5m ahead,
	(0, 0)	Move 5m astern to final position
(6)	(0, 0)	Start position above topside,
	(5, 0)	move 5m to port side,
	(0, 0)	move 5m to starboard to final position
(7)	(0, 0)	Start position above topside,
	(0, -3)	move 3m ahead,
	(3, -3)	move 3m to port side,
	(3, 0)	move 3m astern,
	(0, 0)	move 3m to starboard to final position

different angles. This was done at the same resolution as used in experiment 1 to obtain the most accurate results. The calibration was performed by using the calibration function in OpenCV. For each screen shot the function draws individual chessboard corners detected either as circles if the board was not found, or as colored lines if the board was found as can be seen in Figure 11. This resulted in the camera calibration matrix (eq. 22). The accuracy of the calibration can be indicated by the Root-Mean-Square (RMS) of the reprojection error. The reprojection error is calculated by projecting three-dimensional of chessboard points into the image using the final set of calibration parameters and comparing the position of the corners. An RMS of 0.1 means that, on average, each of these projected points is 0.1 pixels away from its actual position. An acceptable RMS should be as close to zero and at least have a value between 0.1 and 1. The obtained result in equation 23 seems to be accurate enough to continue with experiment 2.

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1258.27 & 0 & 468.746 \\ 0 & 1259 & 296.561 \\ 0 & 0 & 1 \end{bmatrix} \quad (22)$$

$$RMS = 0.0198 \quad (23)$$

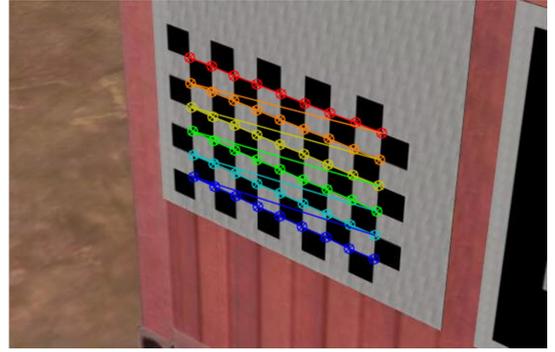


Fig. 11: Camera Calibration process in simulation environment. One of the pictures used to calibrate the camera.

B. Experiment 2: Accuracy and precision

An advantage of testing in a virtual environment is that a log is kept of the historical path of each individual object. All markers are inserted as individual objects with their origin in the center of the marker, the same location used by the algorithm to perform tracking. In this way it is possible to validate the results of the motion tracker. The image resolution used in the experiments was set to 920 x 600 pixels since this resolution resulted in a frame rate of 15 FPS and therefore sufficient measurement accuracy.

1) *Configuration 1:* Procedure 1 and 2 where used to test in what range the algorithm will start working and within what range the results are within the acceptable 0.15m error margin. After these two experiments it can be concluded that the algorithm returns accurate results when the x -offset is within 8 meters and when y -offset is within 12 meters. This is relevant information to know because it tells within what margins the Motion Tracking Algorithm can be used. During procedure 3 and 4 the topside was positioned above the jacket. A realistic Jonswap wave spectrum was applied to generate waves and induce vessel motions. As a result of vessel motions the topside also starts moving. In addition, rain and fog were also simulated. The results for all 4 procedures can be found in table III.

TABLE III: Results configuration 1

Procedure	e_{MAE} [m]	$ e _{MAX}$ [m]	σ_e [m]	μ_e [m]
1 and 2 combined	x: 0.048	x: 0.137	x: 0.048	x: -0.028
	y: 0.043	y: 0.125	y: 0.035	y: -0.035
	z: 0.048	z: 0.104	z: 0.015	z: -0.048
3 and 4 combined	x: 0.032	x: 0.295	x: 0.043	x: -0.019
	y: 0.031	y: 0.252	y: 0.043	y: -0.013
	z: 0.016	z: 0.090	z: 0.019	z: -0.011

The accuracy of procedures 3 and 4 is slightly worse than procedures 1 and 2 but still within acceptable margins. In procedures 3 and 4 the topside showed higher velocities than

during procedures 1 and 2. Considering the 0.7s time delay caused by the running average this will have severe impact on the accuracy.

2) *Configuration 2*: The results from all three procedures are shown in table IV and indicate that the three different procedures give a slightly better result in terms of overall accuracy than configuration 1.

TABLE IV: Results configuration 2

Procedure	e_{MAE} [m]	$ e _{MAX}$ [m]	σ_e [m]	μ_e [m]
5, 6, 7 combined	x: 0.019	x: 0.105	x: 0.022	x: -0.028
	y: 0.023	y: 0.092	y: 0.020	y: -0.035
	z: 0.034	z: 0.082	z: 0.011	z: -0.037

As a verification of the proposed algorithm, several plots comparing the calculated and the logged curves are acquired from the virtual experiments. The resulting x, y , and z -direction performance curves from procedure 5 are given in Figure 12 as an example. It can be seen that the topside moved 5 meters in positive x -direction and then moved 5 meters back in negative x -direction. It's also observable that the topside had an harmonic movement in y -direction. The topside was suspended by the SSCV cranes and was therefore free to move. The height of the topside was kept constant.

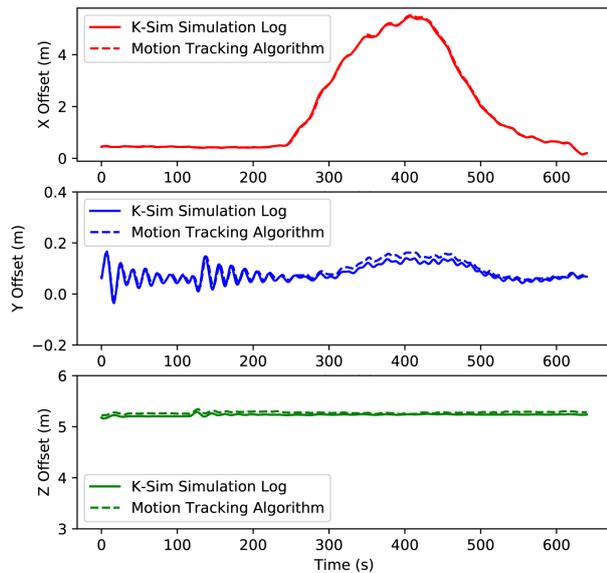


Fig. 12: Combined x, y , and z translations from the simulation log and Motion Tracking Algorithm. Configuration 2, procedure 5.

V. CONCLUSIONS

In this paper, a motion tracking algorithm capable of measuring the relative position between a topside and a jacket is investigated. The proposed method utilized a multiple camera vision system to track the relative motions in real-time. These cameras can be fixed on an SSCV or mounted under a drone. The overall accuracy for all procedures and camera configurations resulted in a maximum standard

deviation of 48 mm. The motion tracking algorithm performed well during experiments, where the accuracy and the precision was tested for different camera configurations. The first configuration used fixed cameras on the stern of the SSCV while the second configuration used drones. Both configurations returned results within the required accuracy but the second configuration - using drones - performed slightly better. A big advantage of drones is the ability to take any desired position with respect to the topside and jacket and are therefore not limited by the view from the SSCV. Since the autonomy level and energy supply of drones are likely to develop in a positive way. Drones could be used for any type of installation in the future, regardless the view from the SSCV. In the meantime drones can be used as a useful visualization tool during topside installations. Drones are able to mimic the presence of people on the jacket and can provide close-ups of every part of the topside or jacket. They eliminate the need for offshore crew to physically access the jacket during installation, which results in a safer work environment.

APPENDIX

The experiment was executed in the Heerema Simulation Center. This simulator runs on the K-SIM platform by Kongsberg with a high-fidelity hydrodynamics and physics engine. State-of-the-art hydrodynamic modelling allows vessels and objects to behave and interact as in real life.

REFERENCES

- [1] Alberto Lopez-Ceron and Jose M. Canas, Accuracy analysis of marker-based 3D visual localization, 2016.
- [2] Garrido-Jurado, S., Muñoz-Salinas, R., Madrid Cuevas, F. J., and Marn-Jimnez, M. J. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 2014. 47(6):22802292.
- [3] Itseez. Open Source Computer Vision Library. 2015. URL <https://github.com/itseez/opencv>, doi:10.1016/j.cell.2011.11.001.
- [4] Vivian W.Y. Tam and Ivan W.H. Fung. Tower crane safety in the construction industry: A Hong Kong study. *Safety Science*, 49(2):208–215, 2011. ISSN 0925-7535. doi: <https://doi.org/10.1016/j.ssci.2010.08.001>.
- [5] Abdulla Al-Kaff, David Martn, Fernando Garca, Arturo de la Escalera, and Jos Mara Armingol. Survey of computer vision algorithms and applications for unmanned aerial vehicles. *Expert Systems with Applications*, 92:447–463, 2018. ISSN 0957-4174. doi: <https://doi.org/10.1016/j.eswa.2017.09.033>.
- [6] Mizuchil, Y., Ogura, T., Kim, YB, et al. J Mech Sci Technol, Camera-based measurement for close-distance relative vessel positioning, (2017) 31: 1899. <https://doi.org/10.1007/s12206-017-0338-3>
- [7] Muñoz-Salinas, R., Marn-Jimnez, M.J., Yeguas-Bolivar, E., & Carnicer, R.M. (2018). Mapping and localization from planar markers. *Pattern Recognition*, 73, 158-171.
- [8] Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer Vision in C++ with the OpenCV Library*. O'Reilly Media, Inc., 2nd edition, 2013. ISBN 1449314651, 9781449314651.
- [9] Konovalenko, I., Kuznetsova, E., Miller, A., Miller, B., Popov, A., Shepelev, D., & Stepanyan, K. (2018). New approaches to the integration of navigation systems for autonomous unmanned vehicles (uav). *Sensors (basel, Switzerland)*, 18(9). doi:10.3390/s18093010
- [10] Lelie, M., Braun, T., Knips, M., Nordmann, H., Ringbeck, F., Zapfen, H., & Sauer, D. (2018). Battery management system hardware concepts: An overview. *Applied Sciences*, 8(4), 534-534. doi:10.3390/app8040534