Vision-based markerless measurement system for relative vessel positioning

Yoshiaki Mizuchi1, Tadashi Ogura1, Youngbok Kim2, Yoshinobu Hagiwara3, Yongwoon Choi1

1Graduate School of Engineering, Soka University, Tokyo, Japan
2Department of Mechanical System Engineering, Pukyong National University, Busan, Korea
3Department of Human and Computer Intelligence, Ritsumeikan University, Shiga, Japan

Abstract: The measurement of relative vessel positioning is important to prevent the risk of vessel collision or the deterioration of working efficiency. This study presents a vision-based system for measuring the relative position and heading between a vessel and a target object without the installation of transponders, reflectors, or markers. Such a markerless measurement system, which is not restricted by the presence or the absence of transponders, is a method of expanding feasible conditions for relative vessel positioning. The proposed system consists of multiple camera units that can measure the distance and direction between an initially designated point on a target object while keep track of the designated point. To evaluate the measurement performance for the proposed system, the authors conducted measurements by using a prototype system in water. The results demonstrated that the system can keep track of the designated points and accurately measure the position and heading of a vessel relative to a target object, despite the vessel’s roll motion.

1 Introduction

A vessel operator is forced to concentrate on the position and heading of a vessel continuously while manoeuvring it during long time relative positioning, such as loading and unloading. The operator’s fatigue caused by this continuous manoeuvring leads to a reduction in working efficiency and safety. In particular, the risk of accidents, such as vessel collision, greatly increases in close proximity to a target. To reduce such risks, the relative position and heading of a vessel must be measured for safer and efficient positioning by triggering an alarm or automatic manoeuvring.

GPS-based techniques have been used as an effective standard approach for measuring a vessel’s position at sea. Ueno [1] proposed a GPS-based system with multiple receivers that provides sufficiently accurate position and attitude of a vessel for berthing operations. Relative positioning will be achieved by measuring both positions of the target and vessel. However, the usable range of centimetre-level GPS-based positioning is limited, because it needs a shore-based reference station for carrier-phase observations. In addition, in close proximity to a target, GPS-based systems risk signal interruptions from tall structures, such as cranes or derricks.

The use of a radar [2] is another dominant technique. This system consists of an interrogator on a moving vessel and several transponders on a target object. The relative position and heading of a vessel are obtained from measured distances and bearings between the sensor and each transponder. Similarly, a laser-based system consisting of a laser sensor on a vessel and reflector on a target also has been used [3]. Although these systems achieve robust and reliable relative position and heading measurements, the usage scenarios of these systems are limited to where transponders or reflectors are placed on a target.

A solution to the problem of measuring the relative position without any installation on a target is markerless measurement using vision sensors. Generally, in vision-based systems, the accuracy of distance measurement relies on the baseline between two cameras. As a result, the proposed system will be inaccurate for the more distant targets. However, the risk of vessel collision decreases with distance from the target. On the other hand, at close proximity, high frequency measurement is necessary for precise positioning of the vessel in order to avoid collision. Hence, vision-based measurement is considered as an optimal solution for vessel positioning at close proximity to the target. A few vision-based systems have been proposed for autonomous navigation [4, 5]. Although these systems efficiently detect and avoid obstacles, they are not designed to measure the position and heading of a vessel relative to a designated target. The authors have proposed vision-based system for relative position and heading measurement [6]. However, this system also needs the installation of specific landmarks for measuring the distance between a vessel and a target [7–9].

In this paper, we propose a vision-based measurement system without any installation on a target object, to expand the applicable range of relative vessel positioning. The proposed vision-based system consists of two camera units with multiple cameras and a motorised camera mount. Each camera unit is capable of keeping track of the designated point on a target object and measuring the distance and bearing to the tracked point. Results of an experiment performed using a prototype system will demonstrate that the proposal is feasible for relative position measurement even during the vessel’s rolling in rough water.

The remaining of this paper is organised as follows. In Section 2, the structure of the proposed system and the principles of position and heading measurements are described in detail. In Section 3, we show the results from our evaluation of the proposed system. Our conclusions are presented in Section 4.

2 Vision-based measurement system for position and heading

2.1 System configuration

Fig. 1 shows the configuration of the proposed position measurement system. This system consists of two camera units placed on a vessel as shown in Fig. 1a. We designate a point on the target object for each camera unit at the start of the measurement. Each camera unit keeps track of the designated point. The vessel’s position (x, y) and heading θ relative to the midpoint of two designated points are given by the distances dx, dy and bearings φx, φy between each camera unit and its designated point.
2.2 Tracking of the designated point

To keep track of the designated point, the positional shift between a reference image and the current input image is calculated successively. The reference image is memorised by the tracking camera, two measurement cameras, and a motorised camera mount, which is called a pan-tilter. The tracking camera has a wide field of view used to track the designated point without losing sight of it. In addition, the tracking camera is used to obtain the bearing to the designated point. The measurement cameras obtain the distance from the camera unit to the designated point. These cameras have a longer focal length than the tracking camera to accurately measure the distance. The pan-tilter rotates the cameras on its pan and tilt axes to keep track of the designated point.

Typically, two types of configurations can be considered for stereo-vision-based measurement with a moving target. One configuration is where actuators rotate each camera, as shown in Fig. 2a. This configuration requires the cross point of the optical axes of the cameras to be on the target for distance measurement. However, the complexity of the system increases for accurate distance measurement because the positional relation between each camera changes continuously. Another configuration is where actuators rotate an entire stereo camera, as shown in Fig. 2b. Considering the need for simple implementation and accurate measurement, we selected the configuration that has a fixed positional relation of cameras.

In practical use, cameras are wobbled continuously by waves or by the motion of the vessel. The positions of the designated point in each camera are temporally shifted by differences in photographic timings. To eliminate positional shift, the photographic timings of all cameras are synchronised by the input of a trigger signal. In the proposed system, we employed a low-distortion lens and a motorised camera mount, which is called a pan-tilter. This configuration requires the cross point of the optical axes, as shown in Fig. 4b. The positional shift between the reference image and the current input image is calculated based on image feature points. In recent years, excellent feature detectors and descriptors have been proposed [10–14]. Although some algorithms [10–12] yield good results in a short time, they have lower accuracy than SIFT (scale invariant feature transform) [13] and SURF (speeded-up robust features) [14]. In this study, we use a GPU-based SURF algorithm, which has sufficient processing speed, scale invariance, and rotation invariance.

Fig. 3a shows detected feature points using a GPU-based SURF algorithm from the reference image. Fig. 3b shows detected feature points from the image captured at a slightly different position. From these detected feature points in each image, point correspondences can be obtained as shown in Fig. 3c. In this figure, each line with a circle represents the displacement between matched feature points. However, these point correspondences include many mismatch pairs, particularly, when the target object has repetitive patterns, such as the images shown in Fig. 3. We eliminate visibly wrong feature matching points by using a voting algorithm based on the Euclidean distances between the matched feature points, as shown in Fig. 3d. Then, the homography between two images is estimated from obtained point correspondences. Here, the mismatch pairs are rejected using the RANSAC algorithm [15]. The arrow in Fig. 3d shows the positional shift of the designated point from the reference image.

The designated point is projected into an input image with a pixel position \((u_o, v_o)\) relative to the optical axis, as shown in Fig. 4a. However, this pixel position is shifted in practical use by the distortion of the lens. The corrected pixel position \((u, v)\) is rescaled according to the following equations

\[
u = u_o \left(1 + k_1 u_o^2 + 2 k_2 u_o^4 + k_3 u_o^6 + 2 p_1 u_o v_o + p_2 (u_o^2 + 2 v_o^2)\right),
\]

\[
v = v_o \left(1 + k_1 v_o^2 + 2 k_2 v_o^4 + k_3 v_o^6 + 2 p_1 u_o v_o + p_2 (u_o^2 + 2 v_o^2)\right),
\]

\[
r^2 = u_o^2 + v_o^2
\]

where \(k_1, k_2, k_3\) and \(p_1, p_2\) denote the coefficients of radial distortion, and \(p_1\) and \(p_2\) denote the coefficients of tangential distortion. Although the distortion factors can be estimated by camera calibration (e.g. [16]), in the proposed system, we employed a low-distortion lens and assumed that all distortion factors are equal to 0 for simplification.

As shown in Fig. 4b, the horizontal angle \(\psi_{\text{ax}}\) between the optical axis and designated point is calculated from the horizontal coordinate \(u\) by using the following equation

\[
\psi_{\text{ax}} = \tan^{-1} \frac{d_x}{f},
\]

where \(f\) denotes the focal length of the lens and \(d_x\) denotes the pixel
size of the imaging device. The vertical angle $\psi_v$ in an image is also calculated in a manner similar to the calculation of the horizontal angle $\psi_h$. By controlling the pan-tilter such that it minimises the obtained angles $\psi_h$ and $\psi_v$, the designated point is continuously tracked.

2.3 Distance and bearing measurement using a camera unit

Fig. 5 illustrates the distance measurement from the camera unit to the designated point on the target. In Fig. 5a, the x-mark denotes the intersection of the target object and another measurement camera’s optical axis. The homography between the two measurement cameras is given in the same manner as the tracking of the designated point. The projected point of the intersection in a camera image given by the homography. The angle $\psi$ between the optical axis and intersection is calculated from (4), by using the intersection instead of the designated point.

The distance $d$ is calculated from the angle $\psi$ using the following equation

$$d = \frac{L_C}{\tan \psi},$$

(5)

where $L_C$ denotes the baseline length between two measurement cameras in Fig. 5a. As shown in Fig. 5b, the distance $d$ between the camera unit and designated point is obtained using the following equation

$$d = \frac{\delta_L + \delta_R}{2},$$

(6)

where $\delta_L$ and $\delta_R$ denote the distances of each measurement camera obtained from (5).

The bearing $\varphi$ to the designated point is determined from the angle $\psi_h$ obtained by the tracking camera and the pan angle $\xi_p$ of the pan-tilter using the following equation

$$\varphi = \psi_h + \xi_p.$$  

(7)

With regard to the vertical angle in an image, the tilt of the camera unit is calculated in a manner similar to the calculation of the horizontal angle $\varphi$.

2.4 Position and heading estimation

Fig. 6 shows the geometry and parameters required for determining the vessel’s position and heading based on the measured distances and bearings. The position and heading of a vessel are calculated from the obtained distances $d_L$, $d_R$ and the bearings $\varphi_L$, $\varphi_R$ of the camera units. The positions of the designated points $(x_{TL}, y_{TL})$ and $(x_{TR}, y_{TR})$ are determined using the following equations

$$x_{TL} = -\frac{L_U}{2} + d_L \sin \varphi_L, \quad y_{TL} = d_L \cos \varphi_L,$$

$$x_{TR} = \frac{L_U}{2} + d_R \sin \varphi_R, \quad y_{TR} = d_R \cos \varphi_R.$$  

(8)

We assume that the length between the camera units $L_U$ is defined in advance, during the installation of camera units. The heading angle $\theta$ for the vessel is also given as

$$\theta = -\tan^{-1}\left(\frac{y_{TR} - y_{TL}}{x_{TR} - x_{TL}}\right).$$  

(9)
The reference point \((x_0, y_0)\) is determined as follows:
\[
x_0 = \frac{x_{TL} + x_{TR}}{2}, \quad y_0 = \frac{y_{TL} + y_{TR}}{2},
\]
In the proposed system, the midpoint between the two designated points is considered the reference point for the vessel’s position. The installed positions of the camera units \((x_{CL}, y_{CL})\) and \((x_{CR}, y_{CR})\) are determined as follows:
\[
x_{CL} = \left(-\frac{L_x}{2} - x_0\right) \cos \theta + y_0 \sin \theta, \\
y_{CL} = \left(-\frac{L_y}{2} - y_0\right) \sin \theta - x_0 \cos \theta, \\
x_{CR} = \left(\frac{L_x}{2} - x_0\right) \cos \theta + y_0 \sin \theta, \\
y_{CR} = \left(\frac{L_y}{2} - y_0\right) \sin \theta - x_0 \cos \theta.
\]
Then, the vessel’s position \((x, y)\) is determined with the following equation:
\[
x = \frac{x_{CL} + x_{CR}}{2}, \quad y = \frac{y_{CL} + y_{CR}}{2}.
\]

### 3 Performance evaluation of the proposed system

To evaluate the measurement performance of the proposed system, we measured the positions and headings of a model vessel by using a prototype system in water. Fig. 7 depicts the experimental setup. The model vessel was 2.0 m long and 1.0 mm wide. The draft (distance between base of the model vessel to water surface) was \(\sim 150\) mm. Table 1 shows the configuration parameters for the prototype system constructed for the experiment. The camera units were mounted \(850\) mm above the base of the model vessel. The baseline length between the camera units was \(1316\) mm. The camera units each had a \(300\)-mm baseline length. The measurement cameras and tracking camera had \(16\)-mm and \(8\)-mm focal lengths, respectively. All cameras had a \(7.5\)-μm pixel size and \(640 \times 480\) image resolution. At the start of the measurement, we designated the points A and B on the wall to each tracking camera as shown in Fig. 7. The distance between the vessel and the designated points was \(\sim 7\) m. To obtain the actual motion of the vessel, we attached a chessboard pattern and IMU (inertial measurement unit, Xsens MTi-10) on the vessel. Here, we obtained the 6 degree-of-freedom motion of the chess board by using a technique for the extrinsic camera parameter estimation [16] with a camera installed on the ceiling above the pool. We assume that the motion of the chessboard is actual motion of the model vessel.

In the experiment, we measured the positions and headings of the vessel under three conditions, wherein the vessel was given \(x\)-translation (surging), \(y\)-translation (swaying), and yawing, respectively. Furthermore, we tested its performance both in the presence and absence of the vessel’s rolling, to evaluate the ability of the proposed system to provide stable measurements in rough water. The amplitude for the roll, which was measured by the IMU, was approximately \(\pm 2.7^\circ\) when the vessel surged, approximately \(\pm 2.5^\circ\) when the vessel swayed, and approximately \(\pm 2.2^\circ\) when the vessel yawed. The period of roll motion was \(1.50–1.72\) s when the vessel surged, \(1.38–1.67\) s when the vessel swayed, and \(1.43–1.53\) s when the vessel yawed.

Figs. 8a, b–10a, b show the measurements of the \(x\) positions, \(y\) positions, and headings of the model vessel, respectively. In each figure, \(a\) and \(b\) show the results without rolling and with rolling, respectively. These positions and headings are the relative positions and headings based on the position and heading at \(0.0\) s. The measured positions and headings using the proposed system are changed according to the vessel’s motion obtained from the chess board. This means that the proposed system can keep track of the designated point and measure the relative positions and headings, even if the vessel is surged, swayed, yawed, and rolled.

### Table 1 Configuration parameters of the prototype system

| Parameter                                         | Value     |
|---------------------------------------------------|-----------|
| baseline length of camera units \((L_u)\)         | 1316 mm   |
| baseline length between measurement cameras \((L_c)\)| 300 mm   |
| focal length of measurement cameras \((L_m)\)      | 16 mm     |
| focal length of tracking camera \((L_t)\)         | 8 mm      |
| pixel size                                        | 7.5 μm    |
| image resolution                                  | 640 × 480 pixel |
Table 2 shows the average error, absolute maximum error, and standard deviation of the relative positions and headings for each condition. The average errors and standard deviations of the $x$ and $y$ positions are less than 0.1 m, which can be assumed sufficient accuracy for relative poisoning, despite the roll motion of the vessel. Similarly, the average errors and standard deviations under each condition are almost the same, irrespective of the presence of roll motion. Although the headings should be measured with 0.1° accuracy for relative vessel positioning, note that the distance between the prototype’s camera units was only 1.3 m. Indeed, this distance would be several magnitudes longer on an actual vessel. This means that the accuracy of headings $\theta$ would improve drastically on an actual vessel. In other words, we can assume that the proposed system would satisfy the sufficient-accuracy requirements of heading measurements on an actual vessel. These results will be inaccurate for more distant targets. However, the measurement accuracy can be satisfied for more distant targets by using a longer baseline between measurement cameras. Thus, the experimental results demonstrate that the proposed system is sufficiently accurate for vessel positioning at close distances, even in waves.

4 Conclusion

In this paper, we proposed a vision-based measurement system for relative vessel positioning without the installation of transponders, reflectors, or markers on a target. The position and heading of a vessel are measured using distances and directional angles between the designated points on the target and the camera units installed on the side of the vessel. The camera unit has the ability to track the designated point using a GPS-based feature point detection and matching algorithm. The distance to the designated point is estimated based on the homography obtained from the feature point correspondences between two measurement cameras. We evaluated the measurement performance of the proposed system given positional displacement and rotation under the conditions of the presence and the absence of roll motion of the vessel in water. The results of this experiment demonstrated that the proposed system can keep track of the designated points and continuously measure the relative positions and headings of a vessel. The
The proposed system has also demonstrated to be sufficiently accurate for relative positioning at close proximity, despite the presence of vessel’s roll motion. From these results, we conclude that the proposed system has the potential to be used as a measurement system for relative positioning, irrespective of whether a target has transponders, reflectors, or visual markers.

5 Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education in Korea (no. NRF-2015R1D1A1A09056885).

6 References

1 Ueno, M.: ‘A GPS-based system for precise shipping guidance and control’, J. Mar. Sci. Technol., 2000, 5, (1), pp. 9–15
2 ‘RADius’, http://www.km.kongsberg.com, accessed 28 November 2015
3 ‘Fanbeam 5’; http://www.renishaw.com/en/fanbeam-25675, accessed 28 November 2015
4 Snyder, F.D., Morris, D.D., Haley, P.H., et al.: ‘Autonomous river navigation’, Proc. of Society of Photo-Optical Instrumentation Engineers, Mobile Robots XVII, 2004, pp. 221–232
5 Huntsberger, T., Aghazarian, H., Howard, A., et al.: ‘Stereo vision-based navigation for autonomous surface vessels’, J. Field Robot., 2011, 28, (1), pp. 3–18
6 Mizuchi, Y., Ogura, T., Kim, Y.B., et al.: ‘Accuracy evaluation of camera-based position and heading measurement system for vessel positioning at a very close distance’. Proc. 15th Int. Conf. on Control, Automation and Systems, 2015, pp. 747–751
7 Mizuchi, Y., Ogura, T., Hagiwara, Y., et al.: ‘Distance measurement system using a stereo camera and radial pattern target for automatic berthing control’, J. Korean Soc. Power Syst. Eng., 2013, 17, (5), pp. 121–127
8 Kawai, H., Choi, Y., Kim, Y.B., et al.: ‘Position measurement of container crane spreader using an image sensor system for anti-sway controllers’. Proc. of Int. Conf. on Control, Automation and Systems, 2008, pp. 683–686
9 Kawai, H., Kim, Y.B., Choi, Y.: ‘Measurement of a container crane spreader under bad weather conditions by image restoration’, IEEE Trans. Instrum. Meas., 2012, 61, (1), pp. 35–42
10 Rosten, E., Drummond, T.: ‘Machine learning for high-speed corner detection’. Proc. of the 2006 European Conf. on Computer Vision, 2006, pp. 430–443
11 Mair, E., Hager, G.D., Burckha, D., et al.: ‘Adaptive and generic corner detection based on the accelerated segment test’. Proc. 2010 European Conf. on Computer Vision, 2010, pp. 183–196
12 Calonder, M., Lepetit, V., Strecha, C., et al.: ‘BRIEF: binary robust independent elementary features’. Proc. 2010 European Conf. on Computer Vision, 2010, pp. 778–792
13 Lowe, D.G.: ‘Distinctive image features from scale-invariant keypoints’, Int. J. Comput. Vis., 2004, 60, (2), pp. 91–110
14 Bay, H., Tuytelaars, T., Van Gool, L.: ‘SURF: speeded-up robust features’, Comput. Vis. Image Underst., 2008, 110, (3), pp. 346–359
15 Fischler, M.A., Bolles, R.C.: ‘Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography’, Commun. ACM, 1981, 24, (6), pp. 381–395
16 Zhang, Z.: ‘A flexible new technique for camera calibration’, IEEE Trans. Pattern Anal. Mach. Intell., 2000, 22, (11), pp. 1330–1334

This is an open access article published by the IET under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0/)