IMPROVING THE LOCALIZATION IN THE DOCKING PROBLEM OF THE VIBOT-2 MOBILE ROBOT

Dinh Quan Nguyen[1,*], Quoc Huy Tran[2]

[1]*Faculty of Aerospace Engineering, Le Quy Don Technical University, Hanoi, Vietnam*

[2]*Tran Dai Nghia University, Ho Chi Minh City, Vietnam*

Abstract

This article presents a positioning solution for a mobile robot system, the Vibot-2 robot, in the docking problem where the robot must move backwards precisely to the charging position or the location to pick up the delivery cart (or return the cart) during the execution tasks. The localization system needs to calculate the robot's posture (including position coordinates and heading angle) in a fixed coordinate system attached to a fixed station with an accuracy within an allowable limit. The solution to build a positioning module for the Vibot-2 robot uses a stereo camera in combination with AprilTags mounted on a locator plate fixed at the station. Compared with the previous positioning solution applied to the Vibot-2 system, the proposed solution in this article changes the camera type and fiducial tag layout. Experimental results show that the new solution helps to increase the field of view of the positioning module by 30%, increase the measurement accuracy by more than 4 times in term of the average error in the horizontal axis, and at the same time, reduce the rate of measurements with errors outside the allowable limit in term of position by more than 1.5 times and in term of heading angle by more than 2 times.

Keywords: *Positioning; AprilTags; mobile robots; docking problem.*

1. Introduction

Nowadays, mobile robots are widely used in manufacturing factories, smart warehouses, performing goods transportation tasks to help businesses improve production efficiency and reduce labor costs. In the medical field, mobile robots are used to transport food, medicine, and necessities to patients [1-3]. In Vietnam, during the Covid-19 pandemic, the Vibot-2 autonomous mobile robot system was used to deliver food to patients in isolation areas, helping to reduce the risk of cross-infection for medical staff [4].

In the transportation problem, an important function for mobile robots is the ability to automatically approach the stations. The docking problems can be automatically returning to the charging station when the robot completes the task or its battery is low. It can also be automatically moving to the location where the delivery

[*] Email: ndquan@lqdtu.edu.vn

DOI: 10.56651/lqdtu.jst.v18.n03.644

carts are located (for example, carts of food, medicine or necessities…) to pick up or return the cart at a specified point. The requirements for positioning problems for mobile robots in these tasks are the need for high accuracy and reliability.

The commonly used positioning solutions for mobile robots mostly use the system of landmarks fixed at the station and combined with observation and measurement devices mounted on the robot. These devices can be Lidar sensors or cameras.

The approaches using Lidars [5, 6] are implemented with a combination of a laser scanner and a specially shaped shield (e.g. the V-shape) for dissection the reference position that the robot needs to reach. This solution is often used in the docking problem where the robot automatically returning to the charging station. But it is not effective in the case of automatically approaching the cart stations because the laser beam is partially blocked by the legs of the cart that causes the loss of input information, thereby reducing the reliability of the positioning measurement.

Using a combination system between cameras and fiducial tags is also a widely applied solution in the docking problem of mobile robots [7-9]. This option has the advantages of low cost, high stability and less affected by changes in lighting conditions. There are many locating techniques using fiducial tags which have been synthesized and compared in [10]. Among them, AprilTags are one of the most commonly used positioning tags in indoor and outdoor applications with the ability to calculate positioning for a fully six degrees of freedom pose. AprilTags were first introduced in 2011 [11] and later enhanced with different libraries, AprilTag 2 [12] and AprilTag 3 [13].

In most robotic applications using cameras and fiducial tags, the localization system consists of only one camera and one fiducial tag (or a bundle of tags) affixed as a reference frame for the mobile robot. Besides, there are rarely any studies that show actual measurements and specific assessments of the stability and accuracy of the localization system when changing the experimental setups (e.g. changing the camera angles and distance between camera and fiducial tag or changing the number of fiducial tags,…). In our previous work [4], we also used a measurement system which consisted of one camera and one fiducial tag bundle of 5 small AprilTags. Although the robot system had succeeded in performing the docking tasks, however, after re-evaluating this solution, we got quite poor results (especially in position estimation) which will be shown later in the following section). Therefore, we propose a solution of using a system of multiple cameras in combination with the AprilTags and give initial results with a stereo camera to evaluate the reliability and accuracy of the localization module under different experiment setups: changing the size and the number of the fiducial tags

(in a tag bundle) as well as changing the relative poses between the stereo camera and the AprilTags-attached frame.

The article is organized as follows: Section 2 briefly presents the mobile robot studied in the article. The pose estimation algorithm is detailed in Section 3. The experimental results are described in Section 4.

2. The Vibot-2 robot and its positioning module in the docking problem

Figure 1 shows the Vibot-2 robot which is an intelligent autonomous mobile robot capable of self-mapping, navigating, and avoiding static and dynamic obstacles. These robots are used in isolation areas with high risk of infection to replace medical staff in transporting food, medicine, necessities, collecting waste and supporting remote medical examination via video call. Each Vibot-2 robot is equipped with appropriate actuators and sensor systems.

The docking problem of the Vibot-2 robot is shown in Fig. 1a and 1b. In both cases, the positioning module uses a fixed plate with a size of 30 × 30 cm with a bundle of five AprilTags which as a reference point (the tag bundle is printed on a white paper with size of 22 × 22 cm, each small tag has the size of 5 × 5 cm), a camera with 90° FOV is centrally mounted on the rear of the robot lower base. The robot will move backwards continuously and will automatically stop when the proximity sensor mounted under the robot rear touches a metal plate installed on the stop barrier at the charging station (or at the cart station).

(a) *(b)*

Fig. 1. The Vibot-2 robot is going backward to a station:
(a) Approaching the charging station; (b) Approaching the food cart pickup location.

The Vibot-2 robot has a structure (2,0), two main driving wheels are fixed wheels mounted on both sides in the middle of the robot base (Fig. 2). The robot moves in term of the linear and angular velocities (V, ω).

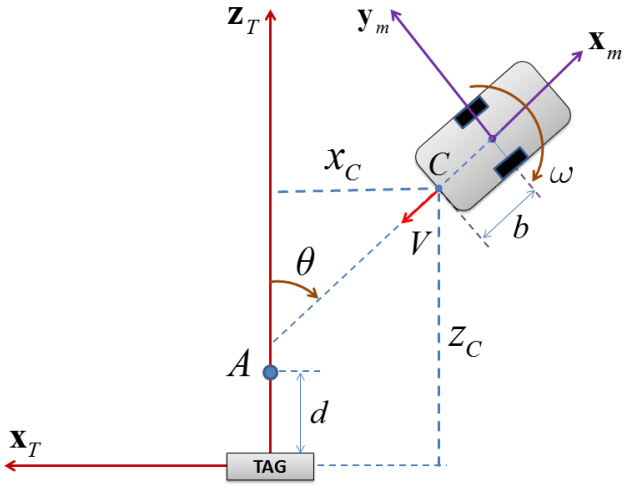

Fig. 2. Robot kinematic model in the docking problem.

In the docking problem, the positioning module calculate the relative pose of a fixed coordinate system attached to the camera center point C with respect to the fixed coordinate system attached to the center of the bundle fiducial tags.

The robot kinematic model can be expressed in term of the coordinates of point C and the robot heading angle as follows:

$$\begin{cases} \dot{x}_C = V\sin(\theta) - b\omega\cos(\theta) \\ \dot{z}_C = V\cos(\theta) + b\omega\sin(\theta) \\ \dot{\theta} = \omega \end{cases} \tag{1}$$

In the docking problem, the linear velocity V has negative values (the robot moves backward toward the fiducial tags). In the ideal case, this process of going backward will be stopped when the robot base moves to a position where we have $(x_C \approx 0, z_C = d, \theta \approx 0)$. In practice, the constraint conditions for the robot's posture in the stationary position need to satisfy the following constraint:

$$\begin{cases} |x_C| \le \Delta_{x\max} \\ z_C = d \\ |\theta| \le \Delta_{\theta\max} \end{cases} \tag{2}$$

These conditions impose a requirement on the accuracy of the position estimator based on the camera system and the fiducial tag with the error to be less than the allowable limit in term of distance $\Delta_{x\max}$ and heading angle $\Delta_{\theta\max}$. Constraint on the z_C coordinate is satisfied by using the proximity sensor mounted under the rear of the robot base. In the case of the Vibot-2 robot, the two limits are $\Delta_{x\max} = 5\,\text{cm}$ and $\Delta_{\theta\max} = 5^{\circ}$.

Fig. 3. Validate the positioning module of Vibot-2 robot.

To evaluate the accuracy and reliability of the applied positioning module of Vibot-2 robot, we carried out an experiment to collect measurement data of (x_C, z_C, θ) at different poses. To change the relative angle between the tag coordinate frame and the coordinate frame attached to the camera center point C, we move the station as shown in Fig. 3 while fixing the the robot base angle. The measurements were performed as follows: at different distances in z-axis, we change the coordinate in x-axis and change also the robot heading angle. Total 157 different poses were measured at $z_d = (0.5, 0.8, 1.2, 1.4, 1.6)\,\text{m}$, $x_d = (-0.8 : 0.2 : 0.8)\,\text{m}$ and $\theta = (-30^o : 10 : 30^o)$. In Table 1, we show the average errors at two close distances $z_d = 0.8\,\text{m}$, $z_d = 0.5\,\text{m}$. We also show the measurement ratios that have the errors falling outside the limit range in term of horizontal distance and heading angle.

Table 1. Invalid measurement ratios and average measurement errors at close distances

Case	**Invalid measurement with** $\lvert e_x \rvert > 5\,\text{cm}$	**Invalid measurement with** $\lvert e_\theta \rvert > 5^o$	**Average errors at** $z_d = 0.8\,\text{m}$	**Average errors at** $z_d = 0.5\,\text{m}$
Applied method	$\frac{87}{157} \approx 55.77\%$	$\frac{45}{157} \approx 28.66\%$	$\bar{e}_x = 11.9242\,\text{cm}$ $\bar{e}_\theta = 1.95^o$	$\bar{e}_x = 7.5108\,\text{cm}$ $\bar{e}_\theta = 1.9246^o$

It can be seen that the applied positioning method has quite good performance in term of heading angle. However, the accuracy in x coordinates is low, much greater than the limit of 5 cm. Besides, the invalid measurement percentages in term of x coordinates as well as heading angle are quite large. Therefore, it is desirable to come up with a new positioning module that can provide better accuracy and reliability.

3. Positioning module using multiple cameras and AprilTags

3.1. Solution using one camera

The positioning module uses the tool package "AprilTags_Ros" [14] to calculate the relative pose between the fixed coordinate system attached to the fixed AprilTag (or

a bundle of tags) with the coordinate system attached to the camera (see Fig. 4).

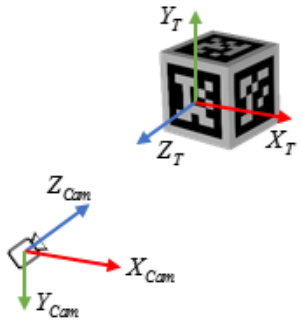

Fig. 4. Localization solution using one camera.

A transformation matrix can be used to represent the pose of the fiducial tag with respect to the camera coordinate frame:

$$ {}^{Cam}\mathbf{T}_T = \begin{bmatrix} {}^{Cam}\mathbf{R}_T & {}^{Cam}\mathbf{P}_T \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \tag{3} $$

where ${}^{Cam}\mathbf{P}_T$ is the position vector and ${}^{Cam}\mathbf{R}_T$ is the rotation matrix. These two terms are computed by the AprilTag_Ros tool.

Since the coordinate frame attached to the AprilTags (the tag-frame) is fixed in space, it is more convenient to represent the coordinates and orientation of the camera with respect to the tag-frame:

$$ {}^{T}\mathbf{T}_{Cam} = \begin{bmatrix} {}^{T}\mathbf{R}_{Cam} & {}^{T}\mathbf{P}_{Cam} \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \tag{4} $$

where

$$ \begin{aligned} {}^{T}\mathbf{R}_{Cam} &= {}^{Cam}\mathbf{R}_T^{-1} \\ {}^{T}\mathbf{P}_{Cam} &= -{}^{Cam}\mathbf{R}_T^{-1}\,{}^{Cam}\mathbf{P}_T \end{aligned} \tag{5} $$

The package Apriltag_Ros use quaternion coordinates to represent the rotation between two coordinate frames. The rotation matrix ${}^{T}\mathbf{R}_{Cam}$ can be expressed in term of a quaternion vector $\mathbf{q} = (q_0, q_1, q_2, q_3)$ [15]. However, it is more convenient to use the Euler angles to represent the rotation between the camera frame and the tag-frame. In this article, we use Euler angles with order of ZYX and the rotation angles can be computed via quaternion coordinates as follows:

$$ \begin{cases} \psi = \text{atan2}\left(2q_1q_2 + 2q_0q_3, 2q_0^2 - 1 + 2q_1^2\right) \\ \theta = -\text{asin}\left(2q_1q_3 - 2q_0q_2\right) \\ \phi = \text{atan2}\left(2q_2q_3 + 2q_0q_1, 2q_0^2 - 1 + 2q_3^2\right) \end{cases} \tag{6} $$

where the angles ϕ, θ, ψ are the angles of the rotations about the X_T, Y_T, Z_T axes, respectively.

Given the condition that the two axes in the Y direction of both coordinate frames are parallel ($Y_{Cam} \parallel Y_T$), the angle θ is also the robot heading angle in the tag-frame.

3.2. Solution using multiple cameras

To increase the reliability of the measuring system, one efficient method is to increase the number of the cameras. Let us consider a system consists of *m* cameras and one locator plate with one or a bundle of AprilTags. These camera can be mount at different positions on the robot base so that they can capture the AprilTags at different angles and cover a larger range.

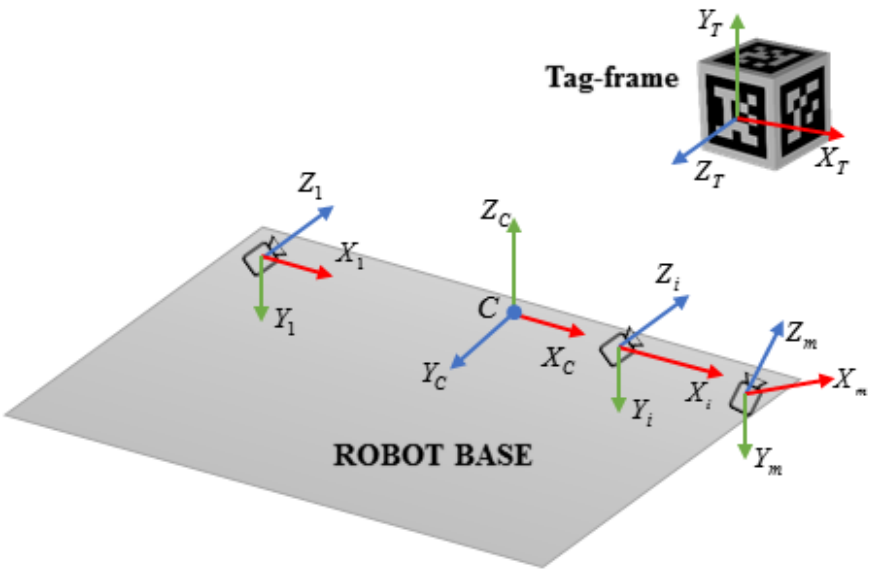

Fig. 5. Localization solution with multiple cameras.

It is necessary to have a reference point C and a reference coordinate system $\langle C \rangle$ attached to C in order to define the robot pose with respect to the tag-frame $\langle T \rangle$. Fig. 5 shows an example of the system.

The *i*-th camera has one coordinate frame $\langle C_i \rangle$ attached to its center point. The transformation matrix that defines the relative pose of the *i*-th camera with respect to the reference coordinate frame $\langle C \rangle$ is known:

$$ {}^{C}\mathbf{T}_i = \begin{bmatrix} {}^{C}\mathbf{R}_i & {}^{C}\mathbf{P}_i \\ \mathbf{0} & 1 \end{bmatrix} \tag{7} $$

Using Apriltag_Ros, one can compute the relative pose between the tag-frame and *i*-th camera frame (i = 1, 2,.., m). In other words, the following transformation matrix is known:

$$ {}^{i}\mathbf{T}_{T} = \begin{bmatrix} {}^{i}\mathbf{R}_{T} & {}^{i}\mathbf{P}_{T} \\ \mathbf{0} & 1 \end{bmatrix} \tag{8} $$

Thus, we have:

$$ \begin{aligned} {}^{T}\mathbf{T}_{C} = {}^{T}\mathbf{T}_{i}\,{}^{i}\mathbf{T}_{C} &= \begin{bmatrix} {}^{i}\mathbf{R}_{T}^{-1} & -{}^{i}\mathbf{R}_{T}^{-1}\,{}^{i}\mathbf{P}_{T} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} {}^{C}\mathbf{R}_{i}^{-1} & -{}^{C}\mathbf{R}_{i}^{-1}\,{}^{C}\mathbf{P}_{i} \\ \mathbf{0} & 1 \end{bmatrix} \\ &= \begin{bmatrix} {}^{T}\mathbf{R}_{C,i} & {}^{T}\mathbf{P}_{C,i} \\ \mathbf{0} & 1 \end{bmatrix} \end{aligned} \tag{9} $$

where

$$ {}^{T}\mathbf{R}_{C,i} = {}^{i}\mathbf{R}_{T}^{-1}\,{}^{C}\mathbf{R}_{i}^{-1} \tag{10} $$

and

$$ {}^{T}\mathbf{P}_{C,i} = -{}^{i}\mathbf{R}_{T}^{-1}\left({}^{C}\mathbf{R}_{i}^{-1}\,{}^{C}\mathbf{P}_{i} + {}^{i}\mathbf{P}_{T}\right) \tag{11} $$

From equations (9) - (11), the relative pose of the tag-frame $\langle T \rangle$ with respect to the reference frame $\langle C \rangle$ are determined from a camera. This results in a set of m possible poses measured from m cameras. Finally, the pose of the robot can be computed as the average sum of this set of poses as follows:

$$ {}^{T}\mathbf{P}_{C} = -\sum_{i=1}^{m} \gamma_{i}\,{}^{i}\mathbf{R}_{T}^{-1}\left({}^{C}\mathbf{R}_{i}^{-1}\,{}^{C}\mathbf{P}_{i} + {}^{i}\mathbf{P}_{T}\right) \tag{12} $$

$$ \theta = \sum_{i=1}^{m} \gamma_{i}\,\theta_{i} \tag{13} $$

where $\gamma_{i}\left(i = 1, 2, ..., m\right)$ is the confident weighing coefficient of the *i*-th camera that satisfies the condition:

$$ \gamma_{1} + \gamma_{2} + ... + \gamma_{m} = 1 \tag{14} $$

and the angle $\theta_{i}\left(i = 1, 2, ..., m\right)$ is computed by applying equation (6) for the rotation matrix ${}^{T}\mathbf{R}_{C,i}$ in (12). Assigning the coefficients to each of the cameras provides us a flexible solution to tune the localization system in practice.

In term of positioning, it is also possible to compute the position vector of the robot in the tag-frame by the means of least square method. By re-writing equation (11):

$$ {}^{i}\mathbf{R}_{T}\ {}^{T}\mathbf{P}_{C,i} = -\left({}^{C}\mathbf{R}_{i}^{-1}\ {}^{C}\mathbf{P}_{i} + {}^{i}\mathbf{P}_{T} \right) \tag{15} $$

we can formulate a system of equations as follows:

$$ \begin{bmatrix} \gamma_1 {}^{1}\mathbf{R}_{T} \\ \gamma_2 {}^{2}\mathbf{R}_{T} \\ \vdots \\ \gamma_m {}^{m}\mathbf{R}_{T} \end{bmatrix} {}^{T}\mathbf{P}_{C} = - \begin{bmatrix} \gamma_1 \left({}^{C}\mathbf{R}_{1}^{-1}\ {}^{C}\mathbf{P}_{1} + {}^{1}\mathbf{P}_{T} \right) \\ \gamma_2 \left({}^{C}\mathbf{R}_{2}^{-1}\ {}^{C}\mathbf{P}_{2} + {}^{2}\mathbf{P}_{T} \right) \\ \vdots \\ \gamma_m \left({}^{C}\mathbf{R}_{m}^{-1}\ {}^{C}\mathbf{P}_{m} + {}^{m}\mathbf{P}_{T} \right) \end{bmatrix} \tag{16} $$

or

$$ \mathbf{M}\ {}^{T}\mathbf{P}_{C} = \mathbf{Q} \tag{17} $$

where

$$ \mathbf{M} = \begin{bmatrix} \gamma_1 {}^{1}\mathbf{R}_{T} \\ \gamma_2 {}^{2}\mathbf{R}_{T} \\ \vdots \\ \gamma_m {}^{m}\mathbf{R}_{T} \end{bmatrix}, \ \mathbf{Q} = - \begin{bmatrix} \gamma_1 \left({}^{C}\mathbf{R}_{1}^{-1}\ {}^{C}\mathbf{P}_{1} + {}^{1}\mathbf{P}_{T} \right) \\ \gamma_2 \left({}^{C}\mathbf{R}_{2}^{-1}\ {}^{C}\mathbf{P}_{2} + {}^{2}\mathbf{P}_{T} \right) \\ \vdots \\ \gamma_m \left({}^{C}\mathbf{R}_{m}^{-1}\ {}^{C}\mathbf{P}_{m} + {}^{m}\mathbf{P}_{T} \right) \end{bmatrix} \tag{18} $$

The solution for (17) is given by:

$$ {}^{T}\mathbf{P}_{C} = \mathbf{M}^{+}\mathbf{Q} \tag{19} $$

with $\mathbf{M}^{+}$ is the pseudo-inverse of matrix $\mathbf{M}$.

3.3. Solution using one stereo camera

In this section, we present the solution of using a stereo camera, which is the Zed2i camera [16] with a large field of view (120° FOV). Fig. 6 shows the new camera which is mounted right above the old camera on the robot rear.

Fig. 6. Camera mounting position on the robot rear.

The stereo camera system is depicted in Fig. 7 with the coordinate frames $\langle L \rangle$ of the left camera and $\langle R \rangle$ of the right camera. Point C is the middle point of the two cameras, dividing the line between the two cameras into two segments with a length of d. The Zed2i stereo camera has a baseline of 12 cm, thus $d = 6\,\text{cm}$.

Since the left camera eye and right camera eye are positioned symetrically with respect to the reference point C, the confident weighing coefficients are selected as $\gamma_L = \gamma_R = \frac{1}{2}$. In this case, equation (12) is chosen to compute the position vector for faster calculation.

The relative pose of the robot with respect to the tag-frame can be computed as follows:

$$ {}^{T}\mathbf{P}_C = -\frac{1}{2}\,{}^{L}\mathbf{R}_T^{-1}\left({}^{C}\mathbf{R}_L^{-1}\,{}^{C}\mathbf{P}_L + {}^{L}\mathbf{P}_T\right) - \frac{1}{2}\,{}^{R}\mathbf{R}_T^{-1}\left({}^{C}\mathbf{R}_R^{-1}\,{}^{C}\mathbf{P}_R + {}^{R}\mathbf{P}_T\right) \tag{20} $$

$$ \theta = \frac{1}{2}\left(\theta_L + \theta_R\right) \tag{21} $$

where subscriptions L, R denote the terms related to the left and right camera eyes respectively and

$$ \begin{gathered} {}^{C}\mathbf{R}_L = {}^{C}\mathbf{R}_R = \mathbf{rot}\left(\mathbf{x}, \pi\right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \\ {}^{C}\mathbf{P}_L = \begin{bmatrix} -d \\ 0 \\ 0 \end{bmatrix},\ {}^{C}\mathbf{P}_R = \begin{bmatrix} d \\ 0 \\ 0 \end{bmatrix} \end{gathered} \tag{22} $$

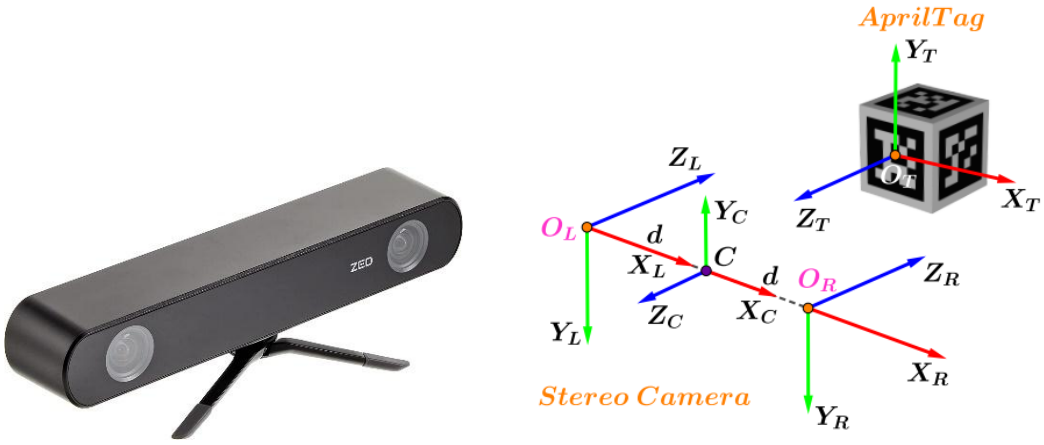

Fig. 7. Coordinate frames of the Zed2i camera and the reference AprilTag.

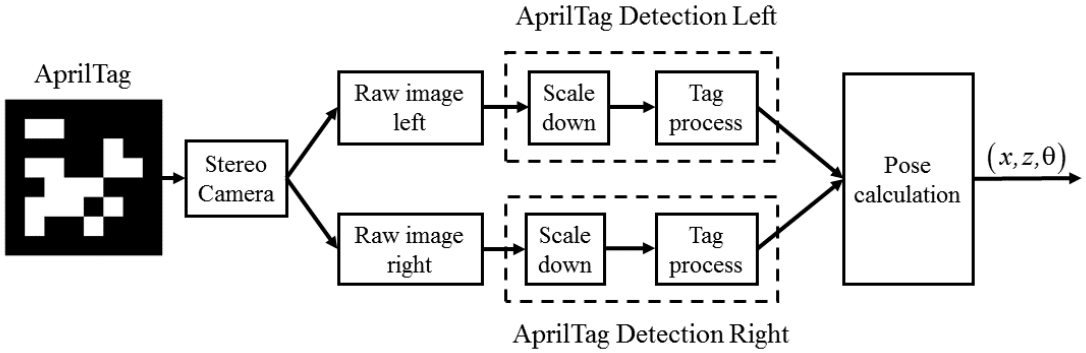

Fig. 8. Calculation of robot pose in the tag coordinate frame.

The rotation matrices ${}^{L}\mathbf{R}_{T}, {}^{R}\mathbf{R}_{T}$ as well as the position vectors ${}^{L}\mathbf{P}_{T}, {}^{R}\mathbf{P}_{T}$ are computed by using Apriltag_Ros tool. The angles θ_L, θ_R are computed from the two rotation matrices ${}^{T}\mathbf{R}_{C,L}$ and ${}^{T}\mathbf{R}_{C,R}$ respectively with:

$$\begin{aligned} {}^{T}\mathbf{R}_{C,L} &= {}^{L}\mathbf{R}_{T}^{-1}\,{}^{C}\mathbf{R}_{L}^{-1} \\ {}^{T}\mathbf{R}_{C,R} &= {}^{R}\mathbf{R}_{T}^{-1}\,{}^{C}\mathbf{R}_{R}^{-1} \end{aligned} \tag{23}$$

The procedure for calculating the robot pose in the tag coordinate frame is shown in Fig. 8, in which the calculations of position of point C - Eq. (20), and heading angle - Eq. (21) are performed in the "Pose calculation" block.

4. Experiment results

To be able to evaluate the efficiency of the new positioning module, measurements were made with the Zed2i stereo camera and Apriltag locator plate at different poses and with different arrangements of the AprilTags on the locator plate. Four measurement cases are carried out with respect to four different arrangements of the AprilTags mounted on the locator plate. Figure 9 shows the setup for the experimental measurements. It is to be seen that the performance of the camera-tag based localization system is depent on:

- The number of fiducial tags in a tag-bundle attached to the locator plate.

- The size of the fiducial tags.

- The relative actual poses (in both position and orientation) between the camera frame and the tag-frame.

Experiments were measured in a uniform light environment, all measuring devices were placed on a table surface with dimensions of 2.2×1.6 m. The AprilTags are printed on a white paper with a maximum size of 30×30 cm (the same size as the

locator plate of the Vibot-2 robot). We fix the locator plate dimension and change the fiducial tag layouts in four different cases.

The size of the AprilTag in case 1 is 9.6×9.6 cm, in case 2 is 15×15 cm and in case 3 is 8×8 cm (a bundle of 3 tags arranged in an equilateral triangle). Case 4 is similar to the case 3 but the tag size is 15×15 cm. The ZED2i stereo camera is mounted on a degree scale and placed in a predetermined position on the table plane, the center of the camera is placed 11 cm below the center of the tag.

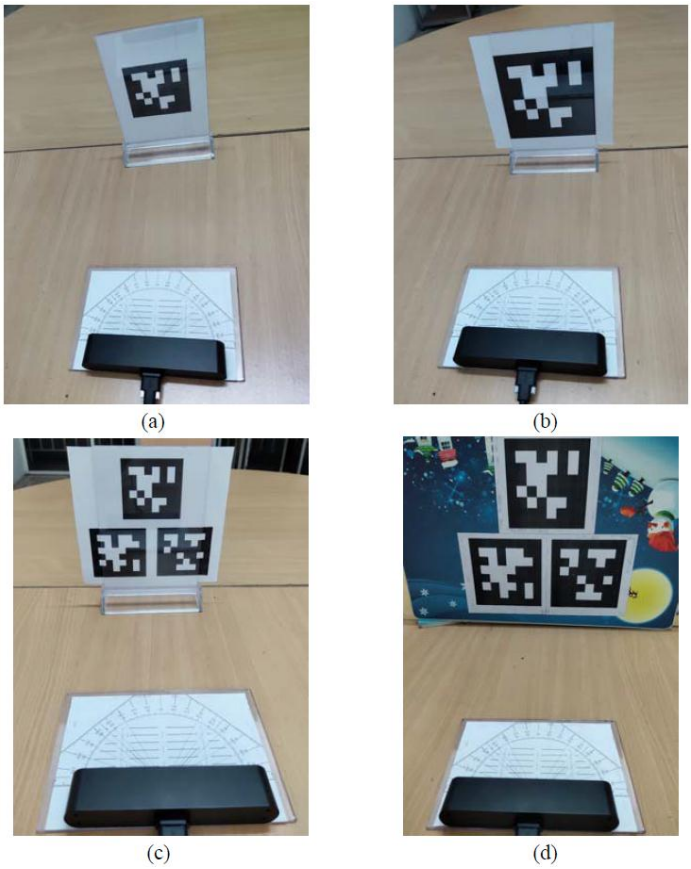

Fig. 9. Experiment setup: (a) Case 1 - tag size 9.6×9.6cm*; (b) Case 2 - tag size* 15×15cm*; (c) Case 3 - bundle of tags, tag size* 8×8cm*; (d) Case 4 - bundle of tags, tag size* 15×15cm.

The AprilTag detection algorithm is run on a computer with ROS Neotic installed. The computer has a CPU Intel i7-10710U 1.1 - 4.7 GHz with 6 Cores, 16GB RAM. The measurement module has an output rate of 25 Hz which is very closed to the output rate of the positioning module that ran on the Vibot-2 robot.

At each experimental location, the camera will be rotated on the degree scale and, at the same time, moved along two coordinate axes on the plane, recording the value in

term of (x, z, θ) (see Fig. 7 for the definition of coordinate systems associated with the camera and the fiducial tag).

The selection of the measurement poses is as follows: Each time the locator plate is measured, it is moved along the z-axis by 0.2 m and in the x-axis by 0.1 m. At each (x, z) coordinates of the locator plate, rotate the stereo camera to angles from -50° to 50° in 10° divisions per rotation. The measurement range in the z-axis is chosen to be (0.4 m ÷ 1.6 m). The selected x-axis measurement range is (-0.5 m ÷ 0.5 m). At each measurement point, the absolute value of the error in x coordinate and direction angle is calculated and displayed on a 3D graph.

The number of measurements in case 1, case 2, case 3 and case 4 are 232, 164, 157 and 150, respectively. Fig. 10 shows the 3D results of measurement absolute errors where the markers "o" represent the valid measurement at each coordinate (x, z), the points marked "x" represent the measurement where the error does not meet the required accuracy of the problem.

Measurement performance statistics are shown in Table 2, which lists information on the percentage of invalid measurement points and average accuracy in the region of close proximity between the camera and the locator plate.

The results in these experiments with the new measuring solution are much better than the old measuring system of Vibot-2 (which are shown in Table 1). In term of positioning, the accuracy has increased by more than 4 times. In term of estimating the heading angles, the accuracy is better at close range between the camera and the locator plate. Furthermore, the new method helps to increase the reliability of the measurement and reduce the invalid measurement percentage by more than 1.5 times in term of position and more than 2 times in term of heading angle.

Table 2. Results of invalid measurement ratio and average accuracy in the region of close proximity

Case	**Invalid measurement with** $\lvert e_x \rvert > 5\text{cm}$	**Invalid measurement with** $\lvert e_\theta \rvert > 5^o$	**Average errors at** $z_d = 0.8\text{m}$	**Average errors at** $z_d = 0.4\text{m}$
1	$\frac{84}{232} \approx 36.2\%$	$\frac{21}{232} \approx 9.05\%$	$\bar{e}_x = 3.4649\text{cm}$ $\bar{e}_\theta = 2.3544^o$	$\bar{e}_x = 2.0442\text{cm}$ $\bar{e}_\theta = 0.9544^o$
2	$\frac{55}{164} \approx 33.53\%$	$\frac{16}{164} \approx 9.75\%$	$\bar{e}_x = 2.8872\text{cm}$ $\bar{e}_\theta = 2.1917^o$	$\bar{e}_x = 1.1958\text{cm}$ $\bar{e}_\theta = 0.7662^o$
3	$\frac{51}{157} \approx 32.48\%$	$\frac{15}{157} \approx 9.55\%$	$\bar{e}_x = 3.0267\text{cm}$ $\bar{e}_\theta = 2.0873^o$	$\bar{e}_x = 1.3947\text{cm}$ $\bar{e}_\theta = 0.8558^o$
4	$\frac{61}{150} \approx 40.67\%$	$\frac{13}{150} \approx 8.67\%$	$\bar{e}_x = 3.3683\text{cm}$ $\bar{e}_\theta = 1.9610^o$	$\bar{e}_x = 1.4941\text{cm}$ $\bar{e}_\theta = 0.7593^o$

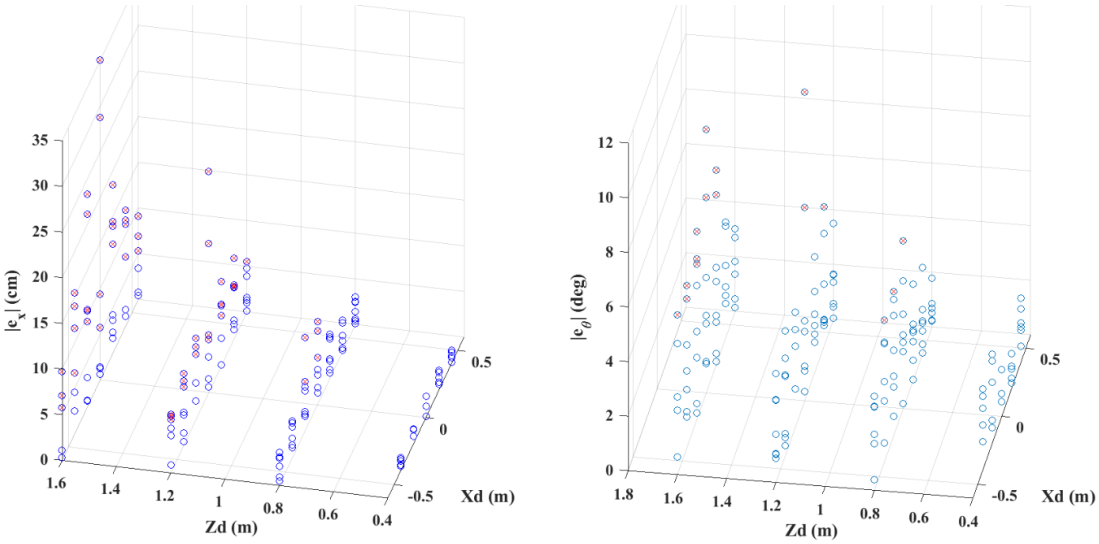

Fig. 10. Experiment results in case 2 with one AprilTag with size 15 × 15 cm.

From the results in Table 2, the tag solution in case 2 gives us the best performance. From these results, we can give some remarks as follows:

- The reliability in measuring heading angle is much higher than the reliability in measuring position in the x-axis (in the 4 cases, the invalid measurement percentage in term of heading angle is less than 10%, meanwhile, the invalid measurement percentage in term of position along x-axis is greater than 33%).

- Increasing the tag size will increase the accuracy (the average error in case 2 with one tag of size $15\times15\text{cm}$ is smaller than case 1 with one tag of size $9.6\times9.6\text{cm}$).

- Changing the number of fiducial tags may change the accuracy; however, from the four cases, it still does not clearly show whether increasing the number of tags will increase the accuracy (the accuracy in case 3 with a bundle of three tags of size $8\times8\text{cm}$ is better than case 1 with only one tag of size $9.6\times9.6\text{cm}$, but the accuracy in case 4 with a bundle of 3 tags of size $15\times15\text{cm}$ is worse than the case 2 with only one tag of size $15\times15\text{cm}$). This means that, in the case of a bundle of tags, the accuracy also greatly depends on the arrangement of the tags.

- At farther distances, the accuracy is reduced. This effect is, in fact, similar to when we increase the tag size.

- The average errors are small in the region of close proximity between the camera and the locator plate. This shows the high reliability of the AprilTag-based positioning solution at closer ranges.

5. Conclusion

The article presented a positioning solution for mobile robots in the docking problem with a system of one stereo camera and AprilTags. The proposed approach can be applied to a general measuring system that consists of several cameras. The initial results in the case of using a stereo camera show great improvements in terms of accuracy and reliability when comparing to the previous applied positioning method on the Vibot-2 mobile robot.

The new localization system presented in this article gives fairly high confidence in term of the average error of the measurements. However, there is still a large proportion of unsatisfactory measurements. Therefore, it is necessary to find a solution to reduce this ratio to improve the reliability of the localization system. Note that in this work, we only change the hardware components which are using a stereo camera and with a different fiducial tag layout (in term of tag number and tag size). Besides, all of the experiments are carried out at static poses. In fact, the failure measurement ratio is greatly increased when the robot is moving.

In future works, we will study the core algorithm of the "Apriltag Detection" tool to improve the quality of input image processing, especially in cases where the camera has unfavorable poses (e.g. at large rotation angles) with respect to the locator plate or when the robot is in motion. At the same time, we also try out different tag layouts by changing the tag number and combinations to further evaluate the effects they have on the performance of the positioning module.

References

[1] J. Evans, B. Krishnamurthy, B. Barrows, T. Skewis, and V. Lumelsky, "Handling real-world motion planning: A hospital transport robot", *IEEE Control Systems*, Vol. 12, Iss. 1, pp. 15-20, 1992. DOI: 10.1109/37.120445

[2] K. Niechwiadowicz and Z. Khan, "Robot Based Logistics System for Hospitals-Survey", IDT Workshop on Interesting Results in Computer Science and Engineering, 2008.

[3] B. Horan, Z. Najdovski, T. Black, S. Nahavandi, and P. Crothers, "OzTug mobile robot for manufacturing transportation", *IEEE International Conference on Systems, Man, and Cybernetics (SMC2011)*. DOI: 10.1109/ICSMC.2011.6084220

[4] V. T. Hoang, Q. N. Tang, X. T. Truong, D. Q. Nguyen, "An indoor localization method for mobile robot using ceiling mounted AprilTags", *Journal of Science and Technique*, Vol. 17, No. 5, pp. 70-91. DOI: 10.56651/lqdtu.jst.v17.n05.531

[5] S. C. Tekkok, B. Bostanci, M. E. Soyunmez, P. O. Ekim, "A novel docking algorithm based on the LiDaR and the V-shape features", *European Journal of Science and Technology*, Special Issue 26, pp. 35-40. DOI: 10.31590/ejosat.947521

[6] S. Vongbunyong, K. Thamrongaphichartkul, N. Worrasittichai, and A. Takutruea, "Automatic precision docking for autonomous mobile robot in hospital logistics – Case-study: Battery charging", *The 11th International Conference on Mechanical Engineering (TSME-ICOME 2020)*. DOI: 10.1088/1757-899X/1137/1/012060

[7] C. S. Sharp, O. Shakernia, and S. S. Sastry, "A vision system for landing an unmanned aerial vehicle", *IEEE International Conference on Robotics and Automation (ICRA2001)*, Vol. 2, pp. 1720-1727. DOI: 10.1109/ROBOT.2001.932859

[8] C. G. Grlj, N. Krznar, and M. Pranjic, "A decade of UAV docking stations: A brief overview of mobile and fixed landing platforms", *Drones*, Vol. 6 (1). DOI: 10.3390/drones6010017

[9] L. A. Mateos, W. Wang, F. Duate, "Bio-inspired Adaptive Latching System for Towing and Guiding Power-less Floating Platforms with Autonomous Robotic Boats", *arXiv*:2001.04293v2. DOI: 10.48550/arXiv.2001.04293

[10] M. Kalaitzakis, Sabrina Carroll, Anand Ambrosi, Camden Whitehead, N. Vitzilaios "Experimental comparison of fiducial markers for pose estimation", *Proceedings of the International Conference on Unmanned Aircraft Systems (ICUAS2020)*. DOI: 10.1109/ICUAS48674.2020.9213977

[11] E. Olson, "AprilTag: A robust and flexible visual fiducial system", *2011 IEEE International Conference on Robotics and Automation (ICRA2011)*. DOI: 10.1109/ICRA.2011.5979561

[12] J. Wang and E. Olson, "AprilTag 2: Efficient and robust fiducial detection", In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2016)*, pp. 4193-4198, 2016. DOI: 10.1109/IROS.2016.7759617

[13] K. Jan, F. Bianca, and W. Hans-Joachim, "Determining and Improving the Localization Accuracy of AprilTag Detection", *IEEE International Conference on Robotics and Automation (ICRA)*. DOI: 10.1109/ICRA40945.2020.9197427

[14] AprilRobotics, "Apriltag_ros", https://github.com/AprilRobotics/apriltag_ros

[15] J. B. Kuipers, "Quaternions and rotation sequences: A primer with applications to orbits, aerospace, and virtual reality", Princeton university press, 1999.

[16] Stereolabs, *ZED 2i – Industrial AI Stereo Camera*. https://www.stereolabs.com/zed-2i/

NÂNG CAO CHẤT LƯỢNG ĐỊNH VỊ TRONG BÀI TOÁN TIẾP CẬN TRẠM DỪNG CỦA HỆ THỐNG RÔ BỐT VIBOT-2

Nguyễn Đình Quân[a], Trần Quốc Huy[b]

[a]*Khoa Hàng không Vũ trụ, Trường Đại học Kỹ thuật Lê Quý Đôn*

[b]*Trường Đại học Trần Đại Nghĩa*

Tóm tắt: Bài báo trình bày một giải pháp định vị cho một hệ thống rô bốt di động, Vibot-2, trong bài toán tự động tiếp cận trạm dừng. Trong đó, rô bốt phải di chuyển lùi chính xác đến vị trí sạc hoặc vị trí lấy xe (hoặc trả xe) chở hàng trong các tác vụ thực thi. Hệ thống định vị cần tính toán tư thế của rô bốt (bao gồm tọa độ vị trí và góc hướng) trong một hệ tọa độ cố định gắn với trạm dừng với độ chính xác nằm trong một giới hạn cho phép. Giải pháp xây dựng mô đun định vị cho rô bốt Vibot-2 sử dụng một stereo camera kết hợp với thẻ định vị AprilTag dán trên một tấm chắn định vị đặt cố định tại trạm dừng. So với phương án đã triển khai trước đây trên hệ thống Vibot-2, giải pháp đề xuất trong bài báo này thay đổi lựa chọn loại camera và kết cấu thẻ định vị. Kết quả thực nghiệm cho thấy giải pháp mới giúp tăng 30% trường nhìn của mô đun định vị, tăng độ chính xác của phép đo hơn 4 lần về sai số trung bình theo trục ngang, đồng thời giảm tỉ lệ phép đo có sai số ngoài giới hạn cho phép về mặt vị trí xuống hơn 1,5 lần và về mặt góc hướng hơn 2 lần.

Từ khóa: *Định vị; AprilTags; rô bốt di động; tiếp cận trạm dừng.*

Received: 24/04/2023; Revised: 11/09/2023; Accepted for publication: 04/12/2023

❑