In: AIAA SciTech Forum 2023, 23-27 January 2023, National Harbor, Maryland, USA
DOI:10.2514/6.2023-2224

Low-cost Multi-object Positioning System with Optical Sensor Fusion

Ravi Ashok Pashchapur*, Yuxi Chen† and Dmitry Ignatyev ‡
Centre for Autonomous & Cyber- Physical System, SATM, Cranfield University, MK43 0AL

Indoor position estimation of any moving objects with the aid of integration of multiple sensors such as optical, radio, and ultrasonic is an ongoing field of research. Although, commercial companies like VICON and OptiTrack provides the higher precision indoor positioning by using custom design optical sensors, but not every teaching or research institute can afford them because of the cost metric. To overcome this problem, an affordable low-cost solution for object tracking using multiple low-cost cameras is provided in this paper. The object tracking system introduced in this paper uses cameras to track active markers such as LEDs and estimate the coordinates of these LEDs with less than 0.15 meters of error in all the axes. The VICON camera system setup is used to opt ground truth measurements which later utilized for error estimation of the low-cost object tracking system.

I. Nomenclature

FOV_x	=	camera horizontal field of view
FOV_y	=	camera vertical field of view
x	=	distance between the two horizontal markers
y	=	distance between the two vertical markers
z	=	distance between the camera and the wall
θ_{cam}	=	camera elevation
ϕ_{cam}	=	camera heading
x_{cam}	=	camera coordinates in X-axis
y_{cam}	=	camera coordinates in Y-axis
z_{cam}	=	camera coordinates in Z-axis
x_{origin}	=	origin coordinates in X-axis
y_{origin}	=	origin coordinates in Y-axis
z_{origin}	=	origin coordinates in Z-axis
x_p	=	coordinates of point generated on line of sight in X-axis
y_p	=	coordinates of point generated on line of sight in Y-axis
z_p	=	coordinates of point generated on line of sight in Z-axis
x_{offset}	=	horizontal offset percentage of target from the center of the screen
y_{offset}	=	vertical offset percentage of target from the center of the screen
θ_{los}	=	elevation of the LOS from camera to target
ϕ_{los}	=	heading of the LOS from camera to target
d	=	distance between camera and the point generated on LOS
J_{int}	=	coordinates for the intersection point of two LOS
err_{int}	=	error of the intersection point
w_{int}	=	weight of the intersection
$J_1 \& J_3$	=	camera coordinates
$J_2 \& J_4$	=	coordinates for the point generated on LOS
$J_a \& J_b$	=	coordinates for the two end points of the shortest line segment of the two LOS

*MSc Autonomous Vehicle Dynamic & Control, Student, SATM.
†MSc Autonomous Vehicle Dynamic & Control, Student, SATM.
‡Senior Research Fellow in Autonomous Systems and Control, SATM..

II. Introduction

In recent years, a position estimation of ground and aerial robotic vehicles in an indoor environment played an important role for autonomous navigation. However, it is quite challenging to perform it in indoor as compared to outdoor positioning. To put it in another word, the Global Navigation Satellite Systems such as GPS, BEIDOU, GLONASS and GALILEO provide coordinates estimation of a mobile object in an outdoor environment using satellites in the orbit, but this system cannot be considered for localization in an indoor arena as the signal from the satellites undergo degradation because of the thick walls and metal surfaces of the building. However, encountering such problems requires the development of robust system. To provide a feasible solution to the above problem, some industries developed motion tracking systems using optical sensors that can track the static as well as moving objects to a millimeter-grade precision. Beside the industrial development, different approaches have been implemented by several researchers to opt a feasible and minimal-cost position estimation system. One of these methods uses an on-board sensor like LIDAR for localization in an indoor environment for GPS-denied navigation [1]. In addition, few other researchers from University of Bristol make use of combination of ultrasonic and radio sensors for indoor positioning in a cross-section area of 8 x 8 meters by decoding radio frequency [2].

Furthermore, only some scholars were very close to an optimal solution, for example, researchers from Universidad Politecnica de Madrid used low-cost optical system and used four distinct colour passive markers to localize aerial vehicle (remote controlled helicopter) in three-dimensional space and this system is only capable of estimating position and orientation within 3 meters [3]. Similarly, another experiment was performed by the researchers from Ecole Militaire Polytechnique, but the problem remains the same, the position was only precise within 2 meters of depth [4]. Hence, finding an optimal method to measure position at greater depth is an on-going challenge. Although this problem has been highly studied in the literature review, but none of the method offered an optimal solution, specifically for multiple objects positions estimation.

In this paper, an indoor positioning system using low-cost cameras and LEDs, capable of providing positions estimation for multiple objects, within at least 10 meters in depth, to a decimeter level of precision, is being developed. The comprehensive design and development of the system are subdivided into three subsections in this paper. Firstly, the section III covers the detailed outline of the localization system and its sub-system; secondly, section IV elaborates the methods and mathematics used for position evaluation; thirdly, followed up by section V where the results are being discussed; and at the last, section VI conclude the paper.

III. System Architecture Overview

In this section the authors of this paper, presents a well-planned structure constructed for the indoor positioning system as shown in Fig 1 step by step. In addition, the detailed hardware and software equipment's utilized in individual subsystems will be discussed. Since this system comprises of multiple sub-systems, a thorough summary of individual subsystems in interior environment is offered.

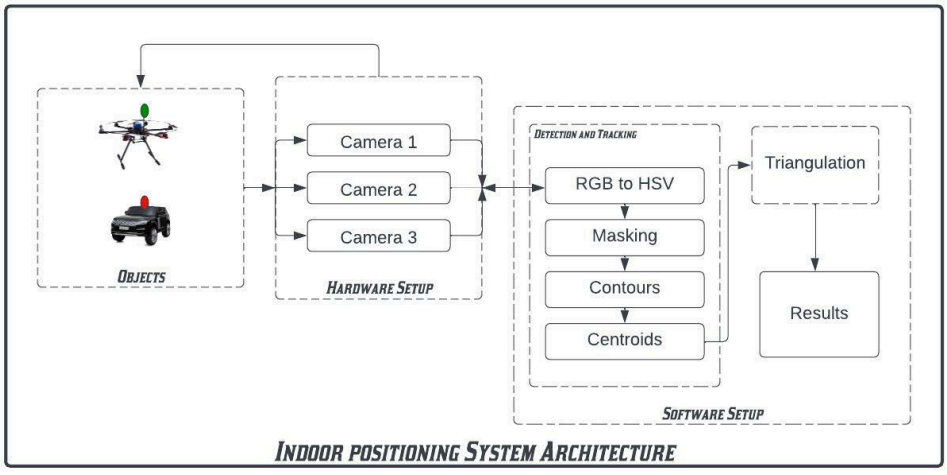


Fig. 1 Representation of indoor system plan.

A. Mobile Objects

The term objects correspond to the dynamic bodies which need to be tracked for indoor positioning experiment. An unmanned aerial vehicle (hexa-copter) and ground vehicle (toy car) as shown in above figure are selected for the experiment as an example. Each vehicle is equipped with a specific colour of LED (active marker) on the top of the vehicle. The hexa-copter drone is integrated with green LEDs and toy car is embedded with red LEDs.

B. Hardware Setup

The hardware sub-system consists of multiple cameras placed in pre-defined location in the arena, and the top view of the environment can be seen in Fig 2. In the experiment, two cameras are used for tracking the LEDs and one camera is placed on the side for observing the experiment. The cameras used for tracking LEDs have the function of adjusting the aperture and focus manually, capable of capturing images with at least 1280 x 720 pixels at 30 frame per second, with horizontal field of view of 36 degrees and vertical field of view of 28 degrees.

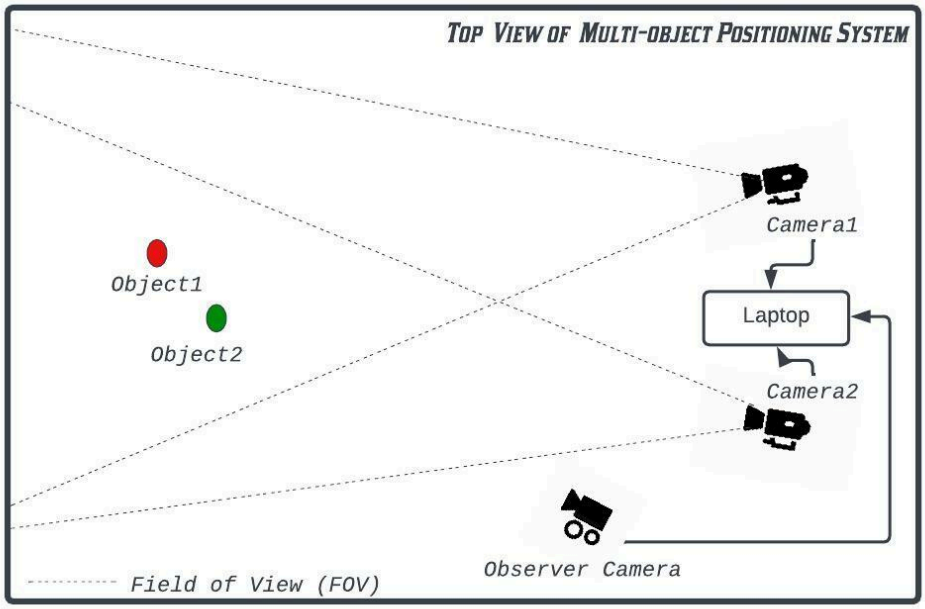


Fig. 2 Top view of experimental arena.

C. Software Setup

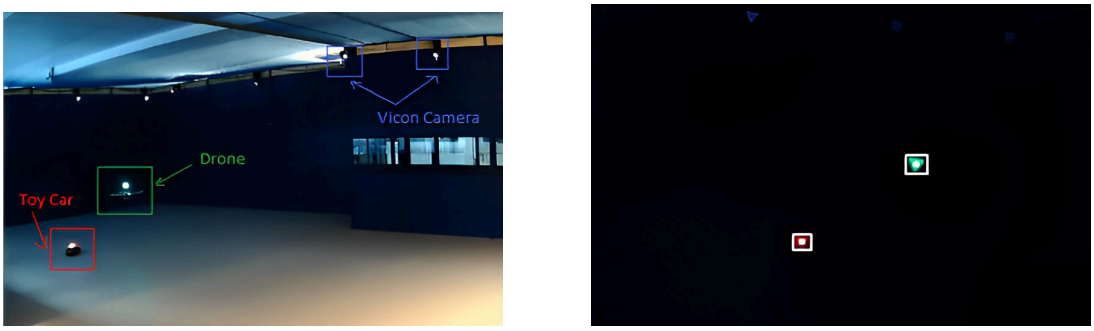
For object detection and tracking sub-system, Open Computer Vision framework is being utilized. In addition, python Integrated Development Environment (IDE) is used for real-time computation and MATLAB is used for plotting and comparing the estimated data with ground truth data.

IV. Design and Development of System

A. Active marker Detection and Tracking

The detection and tracking of colour pixels in image frames are key factors which need to be considered, and it can be done with the help of many different techniques. One of them is by implementing Hue-Saturation-Value (HSV), which is a colour-based method. A researcher from Ecole Militaire Polytechnique [4] used HSV space to detect passive marker (reflective coloured material) but the problem of these markers is when they are more than two metres away from the cameras, cameras start losing track of the markers because of the inconsistent lighting environment. Another approach was by detecting just binary pixels using AR Tags which also encountered the same problem above [5]. To

overcome this problem, a solution is presented where passive markers are completely replaced by active markers and algorithms are introduced in this paper.



(a) *Output from observer Frame* **(b)** *Output from tracking Frame*

Fig. 3 Detection and Tracking of Active Marker.

For multi-object system, LEDs are being used as the active markers, as they emit high intensity light with a wide range of colours and consume very little power comparing to other light sources. For this system, different LED's colours are selected in order to convert the RGB input image frames to HSV image frames for multiple cameras. Detection of the LEDs on the frames are achieved by masking colours with HSV values. Around the colour pixels, a bounding box are created with the aid of contour area as seen in Fig 3b. At the end, the centroid values are computed for each bounding box and fed to triangulation calculation sub-system. The tracking of active marker is done by computing centroids in a loop continuously.

Algorithm 1 Object detection and tracking

```
Initialize cameras and colour pixels;
while True then
    for individual camera in cameras then
        Convert BGR to HSV;
        for individual colour in cameras then
            mask and find contour;
            if contour > Threshold Value then
                draw Bounding Box;
                Calculate Centroid;
            end if
        end for
    end for
    Display frames;
end while
```

B. Position Estimation using Triangulation method

Before setting up all the cameras in the arena, the field of view of the camera is measured and this is done by placing the camera in front of a wall, adjusting the orientation of the camera so that the Line of Sight (LOS) of the camera is perpendicular towards the wall. Furthermore, the distance between the camera and the wall is measured.

Four markers are then put on to the wall so that each marker is at the center point of one of the edges on the capture from the camera which is seen in Fig 4. The distance between the two horizontal markers and the distance between the two vertical markers will be measured. The horizontal field of view and the vertical field of view can then be calculated using below equations [1] and [2].

$$FOV_x = tan^{-1}\left(\frac{0.5x}{z}\right) \times 2 \tag{1}$$

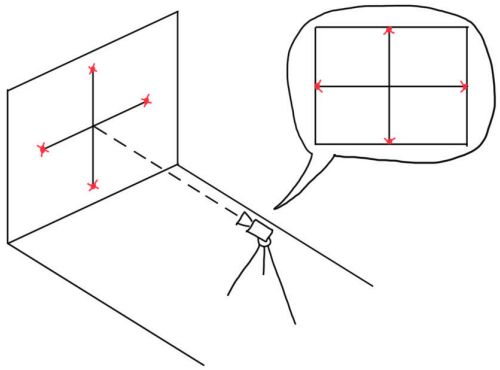


Fig. 4 Estimation of Field of View.

$$FOV_y = tan^{-1}\left(\frac{0.5y}{z}\right) \times 2 \tag{2}$$

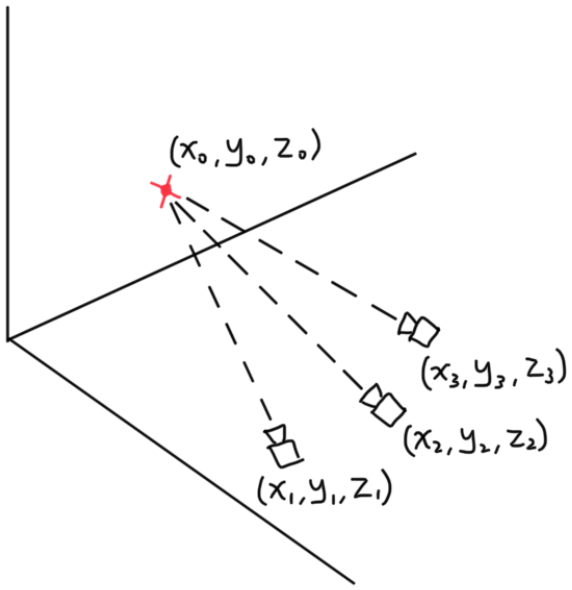


Fig. 5 Calibration Method.

Cameras then need to be setup in the arena. Each camera is placed at its specified location as shown in Fig 5, with the LOS of individual cameras passing through an origin point in which its coordinates are known. The origin point for individual cameras does not need to be the same one if the coordinates of each origin point is known. The heading and elevation of the camera can then be calculated with the following equations [3] and [4].

$$\phi_{cam} = tan^{-1}\left(\frac{x_{origin} - x_{cam}}{y_{origin} - y_{cam}}\right) \tag{3}$$

$$\theta_{cam} = tan^{-1}\left(\frac{z_{origin} - z_{cam}}{\sqrt{(x_{origin} - x_{cam})^2 + (y_{origin} - y_{cam})^2}}\right) \tag{4}$$

When the camera is tracking LEDs, the coordinates on the screen of the camera for individual colours of LED will be saved to an array. These coordinates are then converted to percentages away from the center of the screen, which at the center of the screen, the percentage is 0%, at the left or bottom edge of the screen, the percentage is -100% and at the top or right edge of the screen, the percentage is 100%. The true heading and elevation of the LOS from camera to LED can be calculated with the following equations [5] and [6].

$$\phi_{los} = \phi_{cam} + x_{offset} \times \frac{FOV_x}{2} \tag{5}$$

$$\theta_{los} = \theta_{cam} + y_{offset} \times \frac{FOV_y}{2} \tag{6}$$

An imaginary point then needs to be generated on the LOS for the calculations next step with the following equations [7], [8] and [9].

$$x_p = x_{cam} + d \times \sin(\phi_{los}) \times \cos(\theta_{los}) \tag{7}$$

$$y_p = y_{cam} + d \times \cos(\phi_{los}) \times \cos(\theta_{los}) \tag{8}$$

$$z_p = z_{cam} + d \times \sin(\theta_{los}) \tag{9}$$

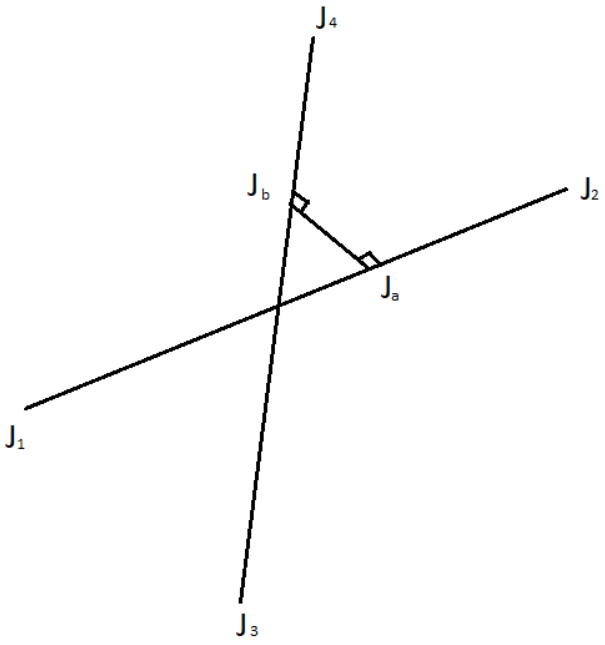


Fig. 6 Representation of pair of line of sights.

In Fig 6, each LOS for one colour of LED will be calculated with all other LOS for the same colour LED to find the intersection points, with the following equations [10] to [19]. In order to compute intersection points, consider two LOS as shown in figure. Usually the two lines in 3D space do not intersect, if they are parallel or if they are coincident but these lines are projected onto plane then shortest line segment can be used as finding intersection points.

To evaluate shortest line segment, this line is defined by two points a and b on lines repsctively. The equations of lines for both the line J_a and J_b is calculated by [10] and [11] respectively.

$$J_a = J_1 + \mu_a \times (J_2 - J_1) \tag{10}$$

$$J_b = J_3 + \mu_b \times (J_4 - J_3) \tag{11}$$

After solving [12] and [13] the two unknown parameter μ_a and μ_b is derived as mention in [15] and [16] and they can be any real number, however, in the developed algorithm it is between 0 and 1 most of the times.

$$d_{1321} + \mu_a \times d_{2121} - \mu_b \times d_{4321} = 0 \tag{12}$$

$$d_{1343} + \mu_a \times d_{4321} - \mu_b \times d_{4343} = 0 \tag{13}$$

$$d_{mnop} = (x_m - x_n) \times (x_o - x_p) + (y_m - y_n) \times (y_o - y_p) + (z_m - z_n) \times (z_o - z_p) \tag{14}$$

$$\mu_a = \frac{d_{1343} \times d_{4321} - d_{1321} \times d_{4343}}{d_{2121} \times d_{4343} - d_{4321} \times d_{4321}} \tag{15}$$

$$\mu_b = \frac{d_{1343} + \mu_a \times d_{4321}}{d_{4343}} \tag{16}$$

The intersection point is calculated by taking the midpoint of shortest line segment [17]. In addition, the intersection weights are computed by taking the inverse of intersection error shown in equation [19].

$$J_{int} = \frac{J_a + J_b}{2} \tag{17}$$

$$err_{int} = \|J_a - J_b\| \tag{18}$$

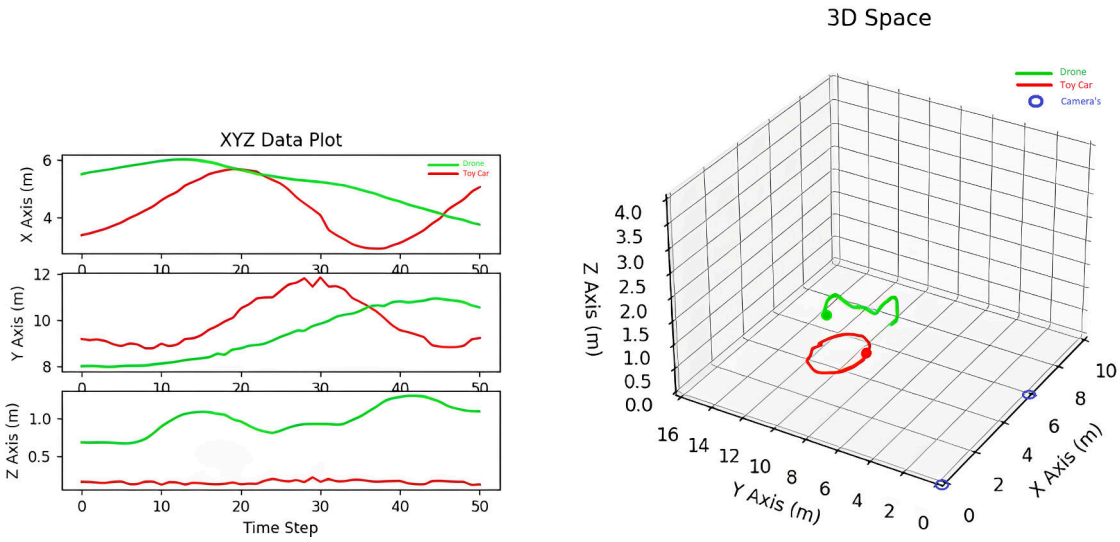
$$w_{int} = \frac{1}{err_{int}} \tag{19}$$

By using the above equations, depending on number of cameras that are tracking the LED, one or more intersection points coordinates, and their corresponding intersection weights will be obtained for that one LED. The estimated coordinates of this LED can be calculated by considering all the intersection points coordinates and weights of this LED using the following equation [20].

$$J_{LED} = \sum \left(J_{int} \times \frac{w_{int}}{\sum w_{int}} \right) \tag{20}$$

V. Simulation Results

To evaluate performance of the proposed positioning system, experiments with mobile testing platforms are performed. The unmanned ground and aerial vehicle equipped with LED markers are operated by two pilots separately within the field of view of both cameras. For simulation, one pilot controlled the ground vehicle in a circular pattern whereas the other pilot controlled the drone above the ground vehicle in random pattern. Meanwhile, the positions of the vehicles estimated by the system are plotted with XYZ axis with respect to time steps, as well as a real time three dimensional space plot, as it is demonstrated below Fig 7a & Fig 7b).



(a) Real-time position data of two mobile object. **(b) Motion of the objects in 3D space**

Fig. 7 Graphical Visualisation of Toy car and Hexa-drone.

When it comes to verification and validation of any system, first thing to focus will be, does it reached the system design requirement or does it achieve the introductory conceptual design. Most of the researchers used statistical methods like Root Mean Square Error (RMSE) of the measurements to estimate the error generated by the localization system. In order to estimate the RMSE of the low-cost positioning system, the VICON camera system is being used as the ground truth. For validation, the reflective VICON passive marker and the LED active marker are mounted onthe drone, and it is being moved around within the field of view of both camera systems. Below Fig 8,9 & 10 show the results of the test.

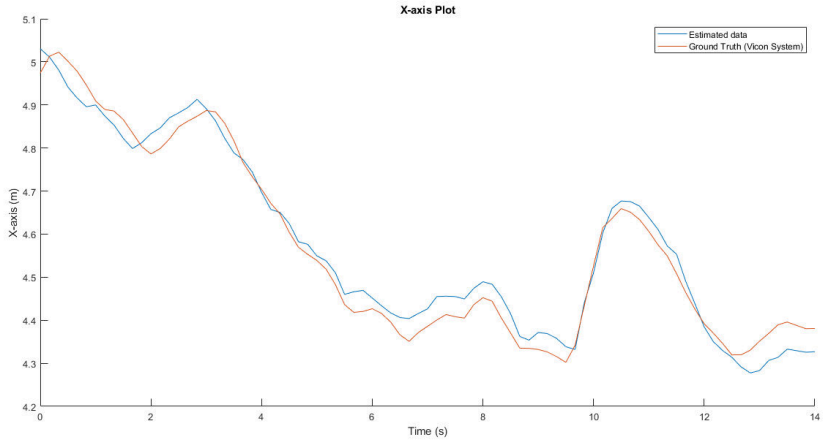


Fig. 8 Relative comparison in X-axis.

As it is showing in the above figures, The data obtained from the low-cost camera system is very close to the data from VICON camera system, with RMSE calculated for XYZ axis to be 0.0336, 0.0482, and 0.0218 meters, respectively. What worth mentioning is that when the reading on Y axis is larger than 10 meters, the error begin to increase. This is because the marker is too far away from the two cameras that even one pixel of error on the frame of the camera will cause a few centimeters of error in position. Coverage of the area can be increased by increasing a number of cameras utilized or by increasing the camera resolution.

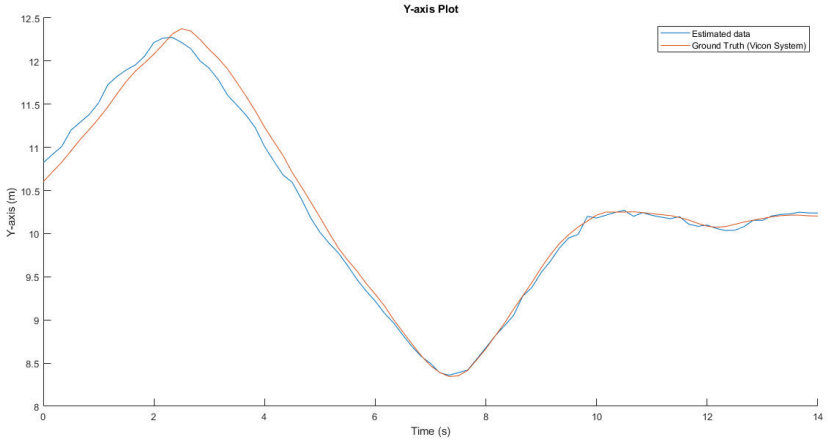


Fig. 9 Relative comparison in Y-axis.

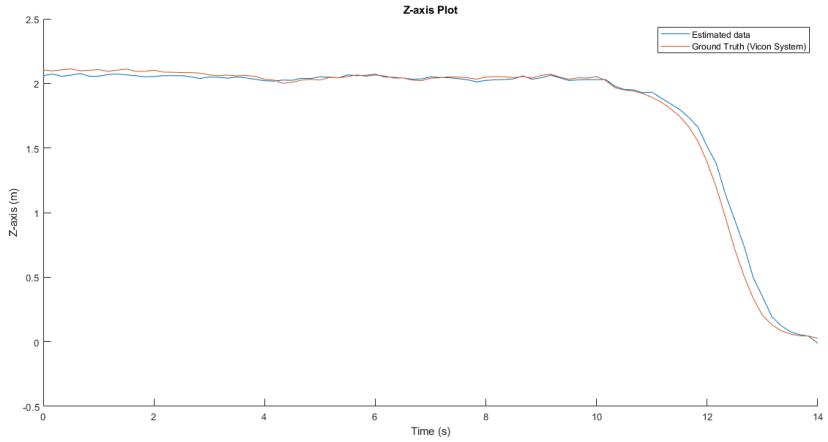


Fig. 10 Relative comparison in Z-axis.

VI. Conclusion

In conclusion, the low-cost multi-object positioning system designed in this paper succeeded in providing relatively high precision position estimation to less than 6 centimetres of error on X and Z axis, and less than 12 centimetres of error on Y axis, for multiple mobile objects that are approximately 10 metres away from the cameras in real time, using some relatively low-cost cameras. The authors of this paper firstly, accomplished the in-detail explanation about the system requirement and its architecture; secondly, the algorithm developed for detection and tracking of active markers; thirdly, the formulas designed for position estimation of multiple mobile object; and at last, the simulation results and comparison with ground truth is performed.

The future work of the project will be improving the system to allow it to estimate the attitude of the vehicle by having multiple different colours of LEDs on one vehicle that the system can calculate the attitude of the vehicle based on the coordinates of these LEDs and feed the position and attitude data back to the controller of the vehicle to allow for a precise indoor autonomous control.

References

[1] Makadia, J., Pashchapur, R., Dhulasawant, T., Gopal G, Y., and Raj P.Y., D., "Autonomous Flight Vehicle Incorporating Artificial Intelligence," 2020. https://doi.org/10.1109/ComPE49325.2020.9200061.

[2] Randell, C., and Muller, H., "Low Cost Indoor Positioning System," 2001. https://doi.org/10.1007/3-540-45427-6_5.

[3] Martínez, C., Campoy, P., Mondragón, I., and Olivares-Méndez, M. A., "Trinocular ground system to control UAVs," 2009. https://doi.org/10.1109/IROS.2009.5354489.

[4] Araar, O., Bouhired, S., Moussiou, S., and Laggoune, A., "Towards Low-Cost Indoor Localisation Using a Multi-camera System," 2019. https://doi.org/10.1007/978-3-030-22750-0_11.

[5] Fiala, M., "ARTag, a fiducial marker system using digital techniques," 2005. https://doi.org/10.1109/CVPR.2005.74.

Acronyms

Augmented Reality (AR), AR Globalnaya Navigazionnaya Sputnikovaya Sistema (GLONASS), GLONASS Global Positioning System (GPS), GPS Hue Saturation Value (HSV), HSV Light Emitting Diode (LED), LED Line of Sight (LOS), LOS Red Green Blue (RGB), RGB Root Mean Sqaure Error (RMSE), RMSE

2023-01-19

Low-cost multi-object positioning system with optical sensor fusion

Pashchapur, Ravi Ashok

AIAA

Pashchapur RA, Pashchapur Y, Ignatyev D. (2023) Low-cost multi-object positioning system with optical sensor fusion. In: AIAA SciTech Forum 2023, 23-27 January 2023, National Harbor, Maryland, USA

https://doi.org/10.2514/6.2023-2224