WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 2, Number 3, Year 2024
DOI: https://doi.org/10.62051/ijcsit.v2n3.09
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Indoor Quadruped Robot Navigation Algorithm Based on ORB-SLAM

Ruoshui Jin[1, 2], Yi Luo[1], Jun Zhao[1, 2]

[1]Sichuan University of Science & Engineering, Zigong, China
[2]Sichuan Key Laboratory of Artificial Intelligence, Yibin, China

ABSTRACT

In recent years, the development of artificial intelligence, big data, and the Internet of Things technologies has revealed unprecedented potential and value in mobile robots across various sectors of automation and intelligence. Among these, quadruped robots have shown unique applications in the domain of indoor security and inspection, as they can navigate multi-storey buildings by ascending and descending stairs. To enhance the stability of visual navigation in indoor inspection environments for quadruped robots, this paper builds on the ORB-SLAM framework. It introduces the AGC algorithm to address the issue of reduced feature point extraction during night patrols, thereby increasing the number of feature points extracted. Additionally, to tackle the problem of high oscillation intensity during the movement of quadruped robots, the EKF algorithm has been incorporated for sensor fusion with IMU, enhancing the robustness of visual navigation. This has been validated through physical experiments.

KEYWORDS

ORB-SLAM, quadruped robot, AGC, EKF

1. INTRODUCTION

With the continuous advancement of technology, mobile robots are increasingly vital in modern society, demonstrating immense potential in enhancing productivity, exploring uncharted territories, and executing high-risk tasks. Bolstered by cutting-edge technologies like artificial intelligence, the Internet of Things, and 5G communication, mobile robots have witnessed significant improvements in autonomy and intelligence. From the precise fertilization of agricultural drones to the intelligent navigation of self-driving cars, and the automated logistics robots in industrial settings, the practical application of these technologies not only showcases the maturity of mobile robot technology but also foreshadows their increasingly critical role in future societies.

Quadrupedal robots, as a significant type of mobile robot, possess distinctive stability and maneuverability, enabling their utilization across a spectrum of complex environments[1]. Unlike wheeled or tracked robots, quadrupedal robots offer enhanced mobility and traversability, maneuvering adeptly across uneven terrain and excelling in extreme conditions like steep slopes, loose mud, and confined spaces[2]. This exceptional mobility, derived from biomimetic limb movement mechanisms, equips quadrupedal robots with remarkable adaptability and exploration capabilities in natural environments.

The rapid progress in robotics, artificial intelligence, and sensor technology has significantly enhanced the autonomy and intelligence of quadrupedal robots[3]. By integrating diverse sensors such as vision systems, inertial measurement units (IMUs), and lidar, quadrupedal robots achieve

high-precision autonomous localization and environment sensing, providing robust technical foundations for executing intricate tasks. Moreover, leveraging advanced algorithms like neural networks and machine learning empowers quadrupedal robots to swiftly and accurately make decisions and responses based on real-time environmental understanding, effectively navigating obstacles and dynamically adjusting action strategies as per task demands[4].

Accurate autonomous localization capability is paramount for mobile robots to enhance their autonomy. The robot's awareness of its position and direction is fundamental for task accomplishment. Localization techniques encompass absolute positioning, determined through GPS or environmental marker identification, and relative positioning, projecting the robot's pose concerning its initial position by measuring motion steps and steering angles[5]. Often, a combination of both methods is employed, with relative localization furnishing short-term precision and absolute localization rectifying accumulated position errors over time[6].

Although existing sensors such as cameras, encoders, and IMUs facilitate position estimation, they are often constrained by specific environmental conditions, lack versatility, and struggle to ensure stable positioning in complex environments[7]. Relying on a single sensor exposes mobile robots to the risk of losing all localization information in case of sensor failure, leading to mission failure or irreparable damage. Therefore, to ensure reliable localization, mobile robots typically integrate multiple sensors and amalgamate the collected data[8].

During task execution, mobile robots must swiftly respond to environmental changes while maintaining motion control stability. Adequate control algorithms facilitate real-time adjustments to the robot's behavior in dynamic environments, enabling it to anticipate and adapt to potential obstacles while executing high-speed tasks, ensuring smooth and safe operation[9].

Further research and enhancements in autonomous localization and motion control of mobile robots will augment their autonomy and enable them to adapt better to unfamiliar environments and execute complex tasks. This expansion of mobile robot capabilities will not only broaden their application scope but also enhance their contributions to human endeavors and daily life. This paper focuses on the scenario of indoor patrol by quadruped robots, emphasizing vision-based localization methods in indoor environments, integrating IMU sensors with vision localization, and exploring quadruped robot path tracking based on localization. The main contributions of this work are surmised as follows:

1. Combining the truncated AGC adaptive gamma correction with the ORB feature point extraction algorithm enables the extraction of a considerable number of feature points at night, thus enabling normal navigation.
2. An observation model for the IMU has been established, and an IMU pose estimation algorithm based on the Extended Kalman Filter (EFK) has been designed.
3. A fusion strategy for the IMU and visual sensors has been devised, and the effectiveness of the fusion algorithm has been validated through simulations and physical experiments.

2. IMPROVED ORB EXTRACTION ALGORITHM

2.1. Problem Formulation

In the ORB-SLAM system, the primary step is to extract feature points from images and perform feature point matching between adjacent frames. Subsequent tasks such as pose transformation, map establishment, and map tracking are all based on the foundation of being able to extract high-quality feature points from images. Therefore, the feature point extraction algorithm is crucial. In practical operation, on one hand, there is a requirement for the quantity of extracted feature points—there should be enough feature points extracted. On the other hand, there is a requirement for speed—it should be fast to extract feature points from images.

In low-light conditions, images captured by cameras may be affected by noise interference, reducing the image quality and making texture features less prominent, thereby reducing the accuracy of feature extraction and matching. In nighttime environments such as hallways, there may be corners where lighting is insufficient or indoor scenes where lights are off. When these scenes appear within the field of view of the visual camera, they can cause significant interference with feature point extraction, leading to a severe reduction in the number of extracted feature points, thereby preventing the SLAM algorithm from continuing to track the current pose. As shown in Figure 1, when the lighting intensity is insufficient, the number of feature points that the ORB algorithm can extract is very limited.

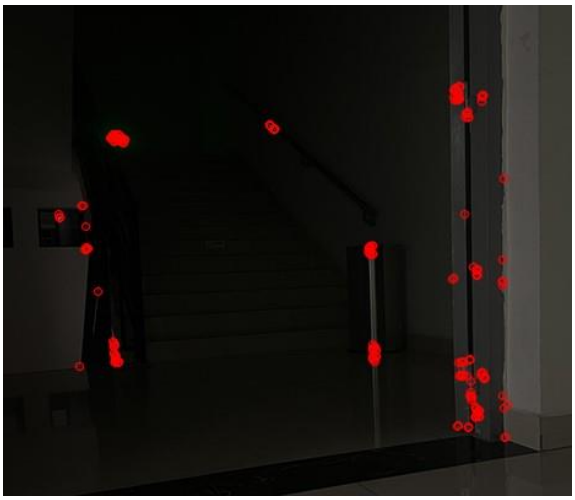


Figure 1. The ORB algorithm extracts feature points from night scenes

2.2. Adaptive gamma transformations

When the lighting is too low, resulting in low image brightness and reduced contrast, object contours become indistinct, affecting the quality of feature point extraction and resulting in a scarcity of extracted feature points. Adaptive Gamma Correction (AGC) is a common algorithm used to adjust the brightness distribution of images. In grayscale images, the standard deviation refers to the standard deviation of the brightness of a point in the image relative to the overall brightness of the image, reflecting the degree of dispersion of pixel brightness in the image. A larger standard deviation indicates greater differences in pixel brightness in the image, resulting in higher contrast, while a smaller standard deviation indicates smaller differences in pixel brightness and relatively lower contrast. The standard deviation of the image can be used to assess the average contrast of the image[10]. If we define the standard deviation of the image as λ, then we can define in the image:

$$I(x,y) = \begin{Bmatrix} ccI_{\text{low}}(x,y) & \lambda \le 0.25 \\ I_{\text{high}}(x,y) & else \end{Bmatrix} \tag{1}$$

The Cumulative Distribution Function (CDF) of the image is:

$$P(\alpha) = \frac{n_\alpha}{N} \tag{2}$$

where α represents the pixel brightness at coordinates (x, y), n_α represents the number of pixels in the image with brightness α, and N represents the total number of pixels in the image. We use the weighted distribution function to smooth the histogram distribution[11]:

$$P_{wd}(\alpha) = P_{\max}\left(\frac{P(\alpha) - P_{\min}}{P_{\max} - P_{\min}}\right)^{i} \tag{3}$$

The total sum of P_{wd} in the image can be calculated as:

$$\sum P_{wd} = \sum_{\alpha=0}^{\alpha_{\max}} P_{wd}(\alpha) \tag{4}$$

The cumulative distribution function of the image can be described as:

$$C_{wd}(\alpha) = \frac{\sum_{\alpha=0}^{\alpha_{\max}} P_{wd}(\alpha)}{\sum P_{wd}} \tag{5}$$

Then the brightness of the image can be described as:

$$\gamma_{wd} = 1 - C_{wd}(\alpha) \tag{6}$$

When running the ORB-SLAM system, the camera sensor needs to maintain an input frequency of over 20Hz. If the image processing speed cannot meet the input frequency, it will lead to decreased localization performance. To improve operating speed and enhance efficiency, one can first judge the brightness of the input image. If the average brightness of the image is too low, further processing can be considered. Let the average brightness of the image be m_I, and define a dim image based on m_I as:

$$I(x,y) = \begin{cases} I_{\text{else}}(x,y) & m_I \geq 80 \\ I_{\text{dark}}(x,y) & m_I < 80 \end{cases} \tag{7}$$

After distinguishing images with low brightness, gamma correction can be used to enhance the performance of under-lit images. To ensure that the adjustment is applied only to the dim areas without affecting parts with normal lighting, one can control the gamma correction parameters using a method of truncating the Cumulative Distribution Function (CDF). Set as:

$$\gamma'_{wd} = \max(\tau, \gamma_{wd}) \tag{8}$$

where τ represents the threshold for truncating the CDF, which can help preserve detailed contour information in the brighter areas. Apply the following transformation to the pixel brightness of the image to be processed:

$$I_{ce}(\alpha) = 255\left(\frac{\alpha}{255}\right)^{\gamma'_{wd}} \tag{9}$$

Figure 2 shows the feature point extraction effect after using AGC (Automatic Gain Control). It can be observed that both the number and distribution of the extracted feature points have significantly improved.

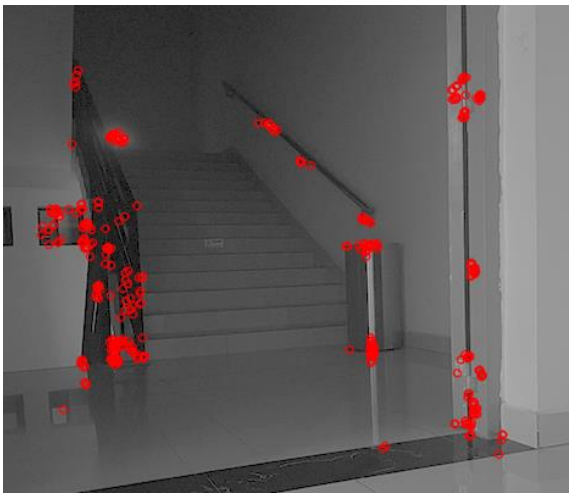


Figure 2. Feature point extraction effect after AGC

3. SENSOR FUSION BASED ON EKF

3.1. Problem Formulation

Robots rely heavily on onboard sensors for positioning while in motion. The higher the quality of the data collected by these sensors, the more accurate the positioning will be. SLAM (Simultaneous Localization and Mapping) algorithms require that sensors maintain a stable and smooth state during movement. Thanks to their legged structure, quadruped robots can perform actions like overcoming obstacles and climbing stairs indoors, which wheeled robots may find challenging. However, the legged structure also induces periodic vibrations during movement, which in turn causes periodic shaking of the sensors mounted on the robot. This vibration can interfere with the SLAM algorithm's positioning calculations, increasing the difficulty of robot localization and reducing the accuracy of pose estimation. This chapter analyzes the problem and combines data from visual and IMU sensors, using sensor fusion techniques to estimate the robot's position.

The quadruped robot can be simplified as four inverted pendulum models. Due to mechanical constraints, as the robot moves, its body vibrates back and forth in the XOZ and YOZ planes, along with the camera sensor fixed on its head, which vibrates periodically. This vibration, whose frequency relates to the current speed and gait, manifests as periodic oscillations. The figure 3 shows the variation over time of the camera sensor's height above the ground as the quadruped robot moves straight ahead.

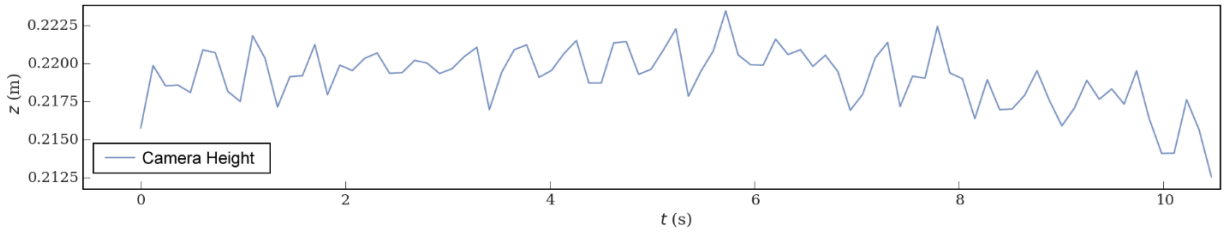


Figure 3. Curve of camera sensor height over time

From the figure 3, it can be seen that the height of the camera sensor continuously fluctuates with an amplitude of about 1 cm, a vibration that persists with the movement of the quadruped robot, leading to inaccuracies in visual positioning. To improve the accuracy of positioning while the quadruped robot is in motion, this chapter establishes an observation model for the IMU and conducts sensor fusion with visual positioning, followed by validation experiments.

3.2. IMU Observation Model

Compared to sensors such as LiDAR and cameras, an Inertial Measurement Unit (IMU) can provide more accurate and stable acceleration and angular velocity information during the motion of a robot. It is low cost, compact size, and easy installation.The IMU is one of the most common measuring units on mobile robots. Comprising an accelerometer and a gyroscope, an IMU provides precise three-dimensional acceleration and angular velocity information. Consider a particle in the coordinate system of the IMU:

$$\boldsymbol{r}_{imu}(t) = (x_1, x_2, x_3)^T \tag{10}$$

To rotate this particle to the world coordinate system, multiply it by a rotation matrix. By taking the derivative with respect to time t, we obtain:

$$\boldsymbol{v}_{world} = R_{WI}\boldsymbol{v}_{imu} + [R_{WI}\boldsymbol{\omega}_{imu}] \times \boldsymbol{r}_{world} = R_{WI}\boldsymbol{v}_{imu} + \boldsymbol{\omega} \times \boldsymbol{r}_{world} \tag{11}$$

where $\boldsymbol{\omega} = R_{WI}\boldsymbol{\omega}_{imu}$, by taking the derivative with respect to time again, we obtain:

$$\begin{aligned} \boldsymbol{a}_{world} &= R_{WI}\dot{\boldsymbol{v}}_{imu} + \dot{R}_{WI}\boldsymbol{v}_{imu} + \boldsymbol{\omega} \times \dot{\boldsymbol{r}}_{world} + \left[\dot{R}_{WI}\boldsymbol{\omega}_{imu} + R_{WI}\dot{\boldsymbol{\omega}}_{imu}\right] \times \boldsymbol{r}_{world} \\ &= \boldsymbol{a} + 2\boldsymbol{\omega} \times \boldsymbol{v} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}_{world}) + \dot{\boldsymbol{\omega}} \times \boldsymbol{r}_{world} \end{aligned} \tag{12}$$

where, $\boldsymbol{a} = R_{WI}\boldsymbol{a}_{imu}$, $\boldsymbol{v} = R_{WI}\boldsymbol{v}_{imu}$.

During measurements, IMUs generate some errors, which mainly consist of deterministic errors (such as biases) and random errors (such as noise). Bias in an IMU refers to a situation where the sensor outputs some non-zero values even when the IMU itself is not subjected to any forces, which can lead to data drift over prolonged periods of motion. Simultaneously, the measurements from an IMU are also susceptible to noise, making the pose information derived directly from integrating IMU data potentially unreliable. Assuming that the bias follows a random walk model and the primary component of the noise is Gaussian white noise, the IMU can be modeled as follows:

$$\begin{aligned} \hat{\boldsymbol{a}}_{imu} &= \boldsymbol{q}_{iw}(\boldsymbol{a}^w + \boldsymbol{g}^w) + b_a + n_a \\ \hat{\boldsymbol{\omega}}_{imu} &= \boldsymbol{\omega}_{imu} + b_\omega + n_\omega \end{aligned} \tag{13}$$

where b_a and b_ω represent the biases, and n_a and n_ω represent the noise components. The noise follows a Gaussian distribution, specifically $n_a \sim N(0, \sigma_a^2)$, $n_b \sim N(0, \sigma_b^2)$.

3.3. IMU Pose Estimation based on EKF

According to the IMU observation model, during the movement of a quadruped robot, we can obtain angular velocities from the gyroscope outputs and accelerations from the three-axis accelerometer outputs. Based on this information, pose changes between two image frames can be calculated by processing the IMU data between them, and subsequently, pose fusion can be performed. Common IMU data processing methods include pre-integration, continuous differentiation, and filtering. The use of the Extended Kalman Filter (EKF) can effectively handle the uncertainties introduced by measurement errors from the gyroscope and accelerometer, minimizing errors caused by drift and noise, optimizing the estimation of system states, and enhancing the accuracy of pose estimation.Assuming a nonlinear Gaussian system, the motion model and observation model are defined as follows:

$$\begin{aligned} \mathbf{x}_k &= g(\mathbf{x}_{k-1}, \mathbf{u}_k) + \boldsymbol{\varepsilon}_k \\ \mathbf{z}_k &= h(\mathbf{x}_k) + \boldsymbol{\delta}_k \end{aligned} \tag{14}$$

where $g(\mathbf{x}_{k-1}, \mathbf{u}_k)$ and $h(\mathbf{x}_k)$ are nonlinear functions, it follows that:

$$\begin{aligned}
\mathbf{x}_{k|k-1} &= g(\mathbf{x}_{k-1|k-1}, \mathbf{u}_k) \\
\mathbf{P}_{k|k-1} &= \mathbf{G}_k \mathbf{P}_{k-1|k-1} \mathbf{G}_k^T + \overline{\mathbf{R}}_k \\
\mathbf{K}_k &= \mathbf{P}_{k|k-1} \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \overline{\mathbf{Q}}_k \right)^{-1} \\
\mathbf{x}_{k|k} &= \mathbf{x}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - h(\mathbf{x}_{k|k-1})) \\
\mathbf{P}_{k|k} &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1}
\end{aligned} \tag{15}$$

where $\mathbf{G}_k$ and $\mathbf{H}_k$ are the partial derivatives of $g(\mathbf{x}_{k-1}, \mathbf{u}_k)$ and $h(\mathbf{x}_k)$ with respect to the system state, respectively:

$$\begin{aligned}
\mathbf{G}_k &= \left. \frac{\partial g(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} \right|_{x = x_{k-1}, u = u_k} \\
\mathbf{H}_k &= \left. \frac{\partial h(\mathbf{x})}{\partial \mathbf{x}} \right|_{x = x_k}
\end{aligned} \tag{16}$$

4. EXPERIMENT

4.1. Experimental Platform

The research platform used in this study is the Unitree Go1 quadruped robot, developed by Unitree Robotics. This model of the robot stands 0.29 meters tall, is 0.58 meters long, and 0.29 meters wide, with a weight of 12 kilograms. It can carry a payload of 3-5 kilograms. It is equipped with two Jetson Nano boards and one Jetson Xavier Nx chip. As far as embedded platforms are concerned, it possesses powerful supercomputing capabilities, enabling real-time visual computing and processing of images received by the camera.

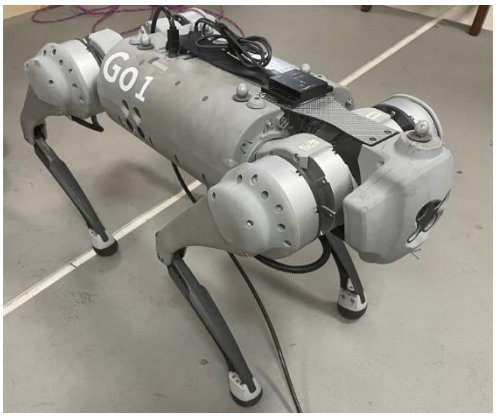
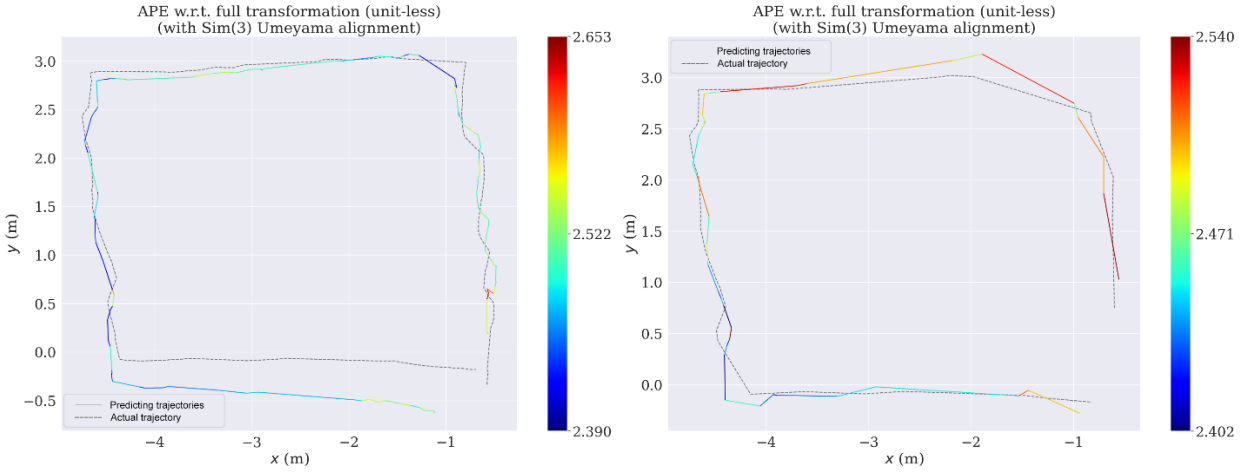


Figure 4. Unitree Go1 Robot

4.2. Quadruped Robot Localization Experiment

For the experiment, a 3x3 meter area within a laboratory was chosen. Using a remote controller, the quadruped robot was navigated around the perimeter of the area to record its path. The tracking trajectory was captured using the OptiTrack motion capture system. Additionally, data from the camera and IMU sensors were recorded into a ROSbag. This ROSbag was then processed using two different localization algorithms: the stereo ORB-SLAM visual algorithm, and a vision+IMU fusion localization algorithm.



(a) Visual positioning trajectory (b) Fusion positioning trajectory

Figure 5. Comparison diagram of visual positioning and fusion positioning trajectories

Table 1. Comparison of positioning errors (centimeters)

Method	Maximum error	Average error	Root mean square error	Standard deviation
Visual	2.6431	2.4663	2.4668	1194.7314
Fusion	2.6534	2.4725	2.4727	262.9235

Trajectory heatmaps for the visual algorithm and the fusion algorithm are respectively shown in Figures 5,the errors shows as Table 1. From the figure, it is evident that incorporating IMU sensor data enhances the robustness of the localization, making the tracked path more closely align with the actual trajectory. Although the maximum and minimum errors are similar, the root mean square error has decreased, indicating smaller fluctuations in localization deviations. This demonstrates the effectiveness of the localization algorithm proposed in this study.

REFERENCES

[1] Chen, Guangrong, et al. "Perturbation-based approximate analytic solutions to an articulated SLIP model for legged robots." Communications in Nonlinear Science and Numerical Simulation 117 (2023): 106943.

[2] Lee, Jin-Woo, Wonjai Lee, and Kyoung-Dae Kim. "An algorithm for local dynamic map generation for safe UAV navigation." Drones 5.3 (2021): 88.

[3] Biswal, Priyaranjan, and Prases K. Mohanty. "Development of quadruped walking robots: A review." Ain Shams Engineering Journal 12.2 (2021): 2017-2031.

[4] Wang, Tianhai, et al. "Applications of machine vision in agricultural robot navigation: A review." Computers and Electronics in Agriculture 198 (2022): 107085.

[5] Fan, Yanan, et al. "A Review of Quadruped Robots: Structure, Control, and Autonomous Motion." Advanced Intelligent Systems (2024): 2300783.

[6] Fusic, S., and T. Sugumari. "A review of perception-based navigation system for autonomous mobile robots." Recent Patents on Engineering 17.6 (2023): 13-22.

[7] Yeong, De Jong, et al. "Sensor and sensor fusion technology in autonomous vehicles: A review." Sensors 21.6 (2021): 2140.

[8] Alatise, Mary B., and Gerhard P. Hancke. "A review on challenges of autonomous mobile robot and sensor fusion methods." IEEE Access 8 (2020): 39830-39846.

[9] Abdulazeez, Adnan Mohsin, and Fayez Saeed Faizi. "Vision-based mobile robot controllers: a scientific review." Turkish Journal of Computer and Mathematics Education (TURCOMAT) 12.6 (2021): 1563-1580.

[10] Rasouli, Amir, and John K. Tsotsos. "The effect of color space selection on detectability and discriminability of colored objects." arXiv preprint arXiv:1702.05421 (2017).

[11] Kim, Mary, and Min Gyo Chung. "Recursively separated and weighted histogram equalization for brightness preservation and contrast enhancement." IEEE Transactions on Consumer Electronics 54.3 (2008): 1389-1397.