Proceedings of the 2025 SICE Festival with Annual Conference
September 9-12, 2025, Chiang Mai, Thailand

Preliminary Study of Automatic Push-Button Operation Using the Quadruped Robot Spot Equipped with an Arm Toward Substation Equipment Maintenance

Toru Kurihara[1†] , Soichiro Hanai[1], Jun Yu[1],Katsuya Hayashi[2], Yoichi Tomiki[2], Kojiro Aoki[2] and Toshihiro Maekawa[2]

[1]School of Information, Kochi University of Technology, Kochi, Japan
(Tel: +81-887-57-2204; E-mail: kurihara.toru@kochi-tech.ac.jp)
[2]Tokyo Electric Power Company Holdings, Inc., Tokyo, Japan

Abstract: With the aging and shrinking workforce responsible for maintaining power equipment, there is an increasing need for research and development of robotic applications in substations to enhance patrol and inspection efficiency, as well as to enable rapid response to equipment failures. In this study, we developed an autonomous push-button operation system using the quadruped robot Spot, which is equipped with a manipulator arm designed for remote operation tasks. Experimental results confirmed that a series of operations —from detecting a 2D tag located beneath the button to executing the button press—can be performed autonomously.

Keywords: Quadruped Robot, Arm Operation, Autonomous Robot, Substation, YOLO

1. INTRODUCTION

With the aging and declining number of personnel responsible for the maintenance of electric power facilities, research and development efforts have been reported to explore the use of robots in substations for improving the efficiency of patrol and inspection tasks[2], as well as for enabling early detection in case of trouble.

While patrol and inspection tasks are expected to be partially automated using commercial products such as Spot and Orbit from Boston Dynamics[1], operations that involve manipulating buttons, valves and switches, or opening and closing control panels remain difficult to automate, although remote operation is possible.

In this study, we investigate the automation of button-pressing operations using a quadruped robot equipped with a manipulator arm, as a step toward fully automating patrol and inspection tasks in substations. The adoption of a quadruped platform is motivated by its ability to traverse uneven and hazardous terrain, where wheeled or tracked robots may be unsuitable. Quadruped robots offer high mobility, obstacle adaptability, and the capability to maintain stability while executing manipulation tasks. Such robots are particularly beneficial for enabling safe and effective operations in environments that are difficult or dangerous for humans — for example, geothermal power plants where volcanic gases are present, or indoor substations where fires have been triggered by equipment failures.

2. MATRIALS AND METHODS

2.1. Quadruped robot Spot®

In this study, we used the quadruped robot Spot developed by Boston Dynamics. The main specifications of Spot are shown in Table 1. The photo of the spot is shown in Fig. 1.

The control program was executed on a laptop PC, which was connected to Spot via Wi-Fi using Spot's access point mode. Commands were transmitted using the SDK provided by Boston Dynamics.

As the target for button operation, we created a mock-up control panel as shown in Fig. 2. The button was mounted at a height of approximately 65 cm from the floor, and an AprilTag was placed at the bottom center of the panel.

Table 1 Main Specifications of Spot® Robot

Item	Specification
Weight	Approx. 32.5 kg
Maximum speed	Approx. 1.6 m/s
Operating time	Approx. 90 minutes (with one battery)
Payload capacity	Up to 14 kg
Slope climbing ability	Up to 30° inclination
Step climbing ability	Up to 30 cm
Ingress protection	IP54
Operating Temprature	-20 °C to 45 °C
Camera configuration	Stereo cameras and depth sensors (front, back, and sides,)
Communication	Wi-Fi, Ethernet

2.2. Overall Processing Flow

The overall processing flow is illustrated in Fig. 3. AprilTag detection and forward/backward movement are implemented by invoking the API provided by Boston Dynamics. When an AprilTag is detected by any of the four onboard cameras mounted on Spot, the robot positions itself approximately 0.7 m in front of the tag, facing it directly. Subsequently, the arm is deployed, and a color camera and depth camera mounted on the gripper

† Toru Kurihara is the presenter of this paper.

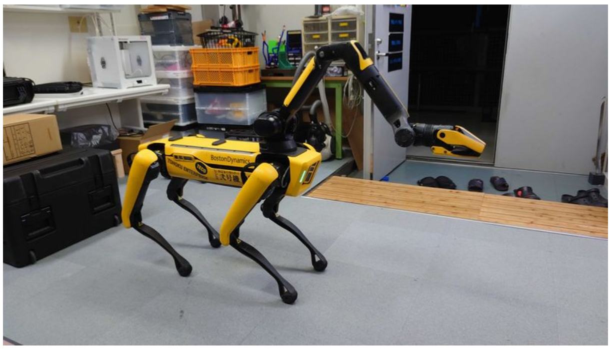


Fig. 1 Spot with manipulation arm.

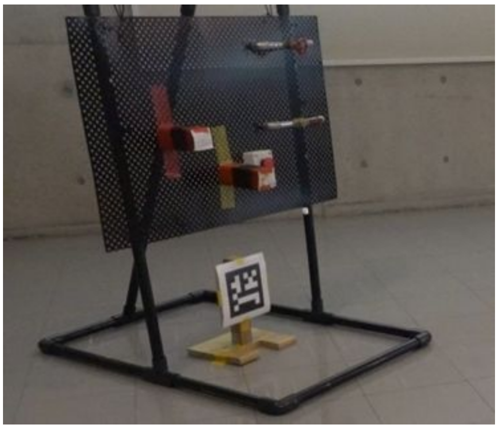


Fig. 2 Panel with target buttton.

are used to capture images. Using the color image, the position of the button is detected via YOLO. Since the Spot API provides aligned color and depth images, the average depth value over a 25 × 25 pixel region centered on the detected button bounding box is extracted from the depth image, and the 3D position in camera coordinate of the button is estimated. Using the force control functionality available in the Spot API, the arm performs a pressing operation with a force of 7 N for 3 seconds, which is appropriate for the mock-up button used in this study. After the pressing operation is complete, the arm is stowed and the program terminates.

2.2.1. Button Detection

The target button used in this study has a diameter of approximately 0.04 m, with a protrusion of about 0.02 m from the surface of the enclosure, as shown in Fig. 4.

To detect the button in color images, we used YOLO11[4] from Ultralytics. Starting from the pretrained YOLOv11n model, we fine-tuned it using eight original button images captured with a compact digital camera, along with mosaic data augmentation, which combines four different images into one during training to improve object detection performance under varied contexts and scales.

2.2.2. Push Button Operation

The Spot API provides several coordinate frames. For example, the odom frame and vision frame represent Spot ' s pose in the global coordinate system. The body

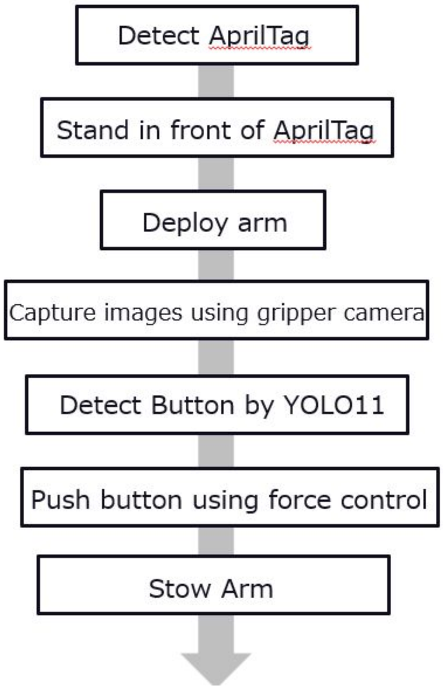


Fig. 3 Overall processing flow.

Fig. 4 Target button size.

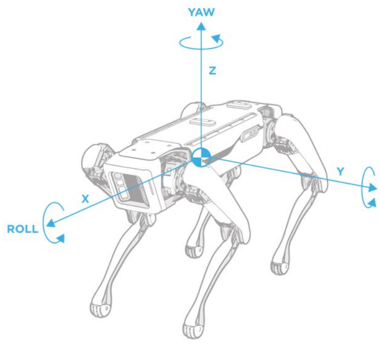


Fig. 5 Spot "body" coordinate from [1]

frame is a coordinate system centered at the body of Spot, with the x-axis pointing forward and the z-axis pointing upward, as illustrated in Fig.5.

Each onboard camera, including the gripper camera, has its own defined camera coordinate frame. Transformations between these frames are continuously updated internally by Spot, and can be retrieved by simply calling the appropriate API functions.

For the button operation, Spot executes the following sequence:

1. The arm is deployed from its stowed position.

2. In the body frame, the arm is moved to coordinates (0.75, 0, 0.05), which corresponds to 75cm forward and 5cm upward from Spot ’ s center.
3. Color and depth images are captured using the onboard cameras.
4. The color image is processed with YOLO11 to detect the center position of the button. If the button is not detected, the system repeats image acquisition.
5. The center position of the button is obtained from the depth image, yielding coordinates in the camera frame.
6. The button position is then transformed from the camera coordinate frame to the body frame.
7. Since the button is located directly in front of Spot, the arm is moved to a position 0.02m in front of the detected button position along the x-axis.
8. The arm is then pushed forward with a force of 7N along the x-axis to press the button.
9. After a 3-second hold, the arm returns to the coordinates (0.75, 0, 0.05).
10. The arm is then retracted to its stowed position.

3. RESULTS

Both the color and depth images captured by the gripper camera are shown in Fig. 6. In the depth image, the light blue text indicates the distance from the camera. Since the panel is made of a perforated metal sheet, there are many areas where the depth camera cannot accurately estimate distances. The button box appears in the lower left of Fig.6(b), and the protrusion of the button can be confirmed: the distance to the button top is 0.355m, compared to 0.380m for the button box surface. Based on this depth information, the arm can be accurately positioned in front of the button.

Next, we confirmed the automatic push-button operation. Spot was randomly placed in front of the mock-up panel, and in most trials, it successfully performed the button press. An example of the operation sequence is shown in Fig. 7.

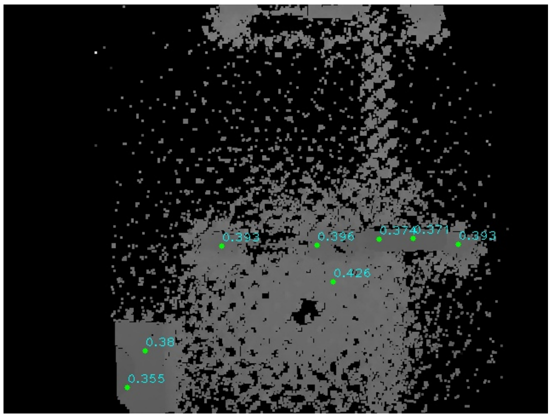
In the experiment, Spot autonomously detected the AprilTag located below the mock-up panel and moved to a position in front of it. It then deployed its arm and captured images. The button was detected using YOLO11, and the arm was moved to the button’s location. Spot pressed the button for approximately three seconds and then retracted the arm.

Occasionally, Spot failed to stop precisely in front of the AprilTag. When the position error threshold was set too small, Spot spent a long time adjusting its position. Conversely, when the threshold was too large, the gripper camera sometimes failed to capture the button correctly.

4. DISCUSSION

A notable limitation identified in this study is the difficulty of precise positioning relative to AprilTags. Although the robot generally accepts a positional error margin of up to 10 cm, it sometimes continues to perform fine adjustments for an extended period without achieving

(a) Color Image.



(b) Depth Image

Fig. 6 Captured image by gripper camera.

successful alignment. We believe this behavior is not specific to the Spot robot, but rather reflects a broader challenge inherent to quadruped robotics. Unlike wheeled platforms, quadrupeds rely on foot placement and full-body coordination for fine movements, which can introduce delays and instability when high precision is required.

However, Spot has the capability to make small positional and orientation adjustments by shifting its body posture without moving its feet. Leveraging this functionality, a more efficient strategy may involve using legged locomotion to reach a coarse target zone, followed by fine-tuning the body position through posture control to bring the manipulator arm into effective range of the target (e.g., a button). Future work will focus on integrating this hybrid approach of locomotion and posture-based refinement to improve task efficiency and stability.

5. CONCLUSION

This study aimed to contribute to the future automation of substation operations by developing an automated button-pressing task using the quadruped robot Spot equipped with a manipulator arm. Through experiments using a mock-up control panel, we confirmed the feasibility of automating such physical interaction tasks. These results demonstrate the potential of legged robotic systems with arms in substation environments, particu-

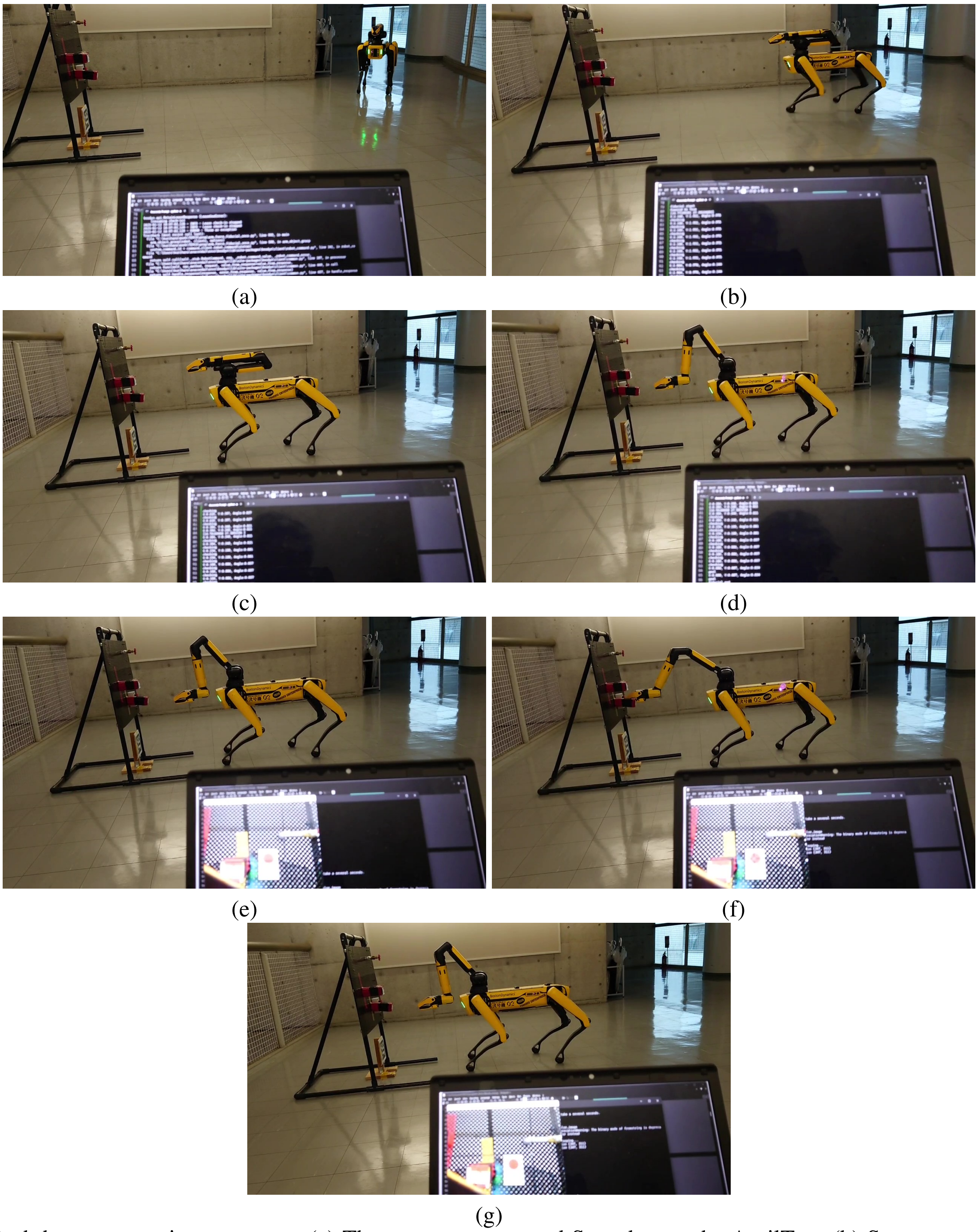


Fig. 7 Push-button operation sequence: (a) The program starts and Spot detects the AprilTag. (b) Spot approaches the AprilTag. (c) Spot stops in front of the AprilTag. (d) Spot deploys its arm. (e) Spot captures images. (f) Spot presses the button. (g) Spot returns the arm to its original position. .

larly for tasks that require both mobility and manipulation.

As future work, we plan to improve positioning accuracy by combining coarse footstep-based locomotion with fine posture adjustments, enabling the arm to reach targets more reliably without prolonged alignment time near fiducial markers.

Furthermore, we plan to extend our approach to automate additional tasks such as valve and lever operations, rotation of rotary selector switches, and opening/closing of control panels.

REFERENCES

[1] " Spot® Boston Dynamics. " [Online]. Available: https://www.bostondynamics.com/spot

[2] Yoichi Tomiki, Katsuya Hayashi, Kojiro Aoki, Toshihiro Maekawa, Shinjou Iida, Kiyomitsu Takeda and Takuya Yokota, "Evaluation of the use of quadruped robot in substations," National Convention record I.E.E. Japan,, 2025.

[3] Simon Zimmermann, Roi Poranne and Stelian Coros, "Go Fetch! - Dynamic Grasps using Boston Dynamics Spot with External Robotic Arm," IEEE In-

ternational Conference on Robotics and Automation (ICRA), pp.4488–4494 , 2021.

[4] Glenn Jocher and Jing Qiu, "Ultralytics YOLO11," [Online] Available: https://github.com/ultralytics/ultralytics