nature portfolio

Peer Review File

Omnidirectional Motion of an Untethered Tripodal Microrobot Using Radial Piezoelectric Actuators

Corresponding Author: Professor Yingxiang Liu

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This manuscript developed a novel miniature tripodal piezoelectric robot that demonstrates omnidirectional locomotion with a zero-radius turning capability. Notably, the design enables smooth transitions across a wide velocity range—from micrometer to centimeter scales—showcasing agile navigation in macroscopic environments and precise motion control in microscopic scenarios. The robot's circular symmetric architecture further allows compact, stackable integration of multiple functional modules. The authors also present closed-loop trajectory control and multi-robot cooperation capabilities. Overall, this work represents an advancement in structural design, steering performance, and system scalability of piezoelectric miniature robots. The idea and concept of the paper are excellent and novel, but the available data on this particular aspect appears to be insufficient. The authors are encouraged to consider the following suggestions for enhancing the manuscript in preparation for future submissions.
Major Comments
(1)The authors claim that the robot exhibits exciting zero-radius omnidirectional turning capability. However, the experiments only demonstrate several specific turning angles. It is necessary to provide a complete set of turning performance data, covering rotation angles from 0° to 360° (with a suitable step size, e.g., 5°), to substantiate the claim.
(2)The environmental adaptability of the miniature robot is also of great concern. Could the authors evaluate its performance in more challenging terrains, such as uneven surfaces, slopes, or gravel grounds? Furthermore, how well does the robot maintain its locomotion capability after compression or drop impacts?
(3)The demonstrated wide velocity range—from 3 cm/s to 5.63 μm/s—is impressive, showing potential for both macro- and micro-scale explorations. However, I am particularly concerned about the precision of velocity control, which directly determines the smoothness of task execution.
(4)From Fig. 3c and 3d, the velocity difference between the two-leg and one-leg motion modes appears marginal. Can I understand that the two-leg motion consumes approximately twice the power of the one-leg mode? If so, does this imply that the robot's energy efficiency decreases in certain motion directions?
(5)The section titled Multi-sensor Fusion Application may be slightly overstated. The experiment primarily demonstrates closed-loop trajectory tracking using a hand-drawn path, relying on an external camera and an onboard attitude sensor—only two sensors without redundancy. A more accurate title or clarification of the "fusion" scope is recommended.
Minor Comments
(6)The robot's endurance of 20 minutes is reported, but it is difficult to assess whether this represents an advantage. It may be more convincing to report the total travel distance achievable under full battery conditions.
(7)The circularly symmetric design allows potential modular stacking, which is interesting. However, additional stacking would increase load. What is the load capacity or payload limit of the robot?
(8)Since miniaturization and a small turning radius are critical for operations in confined spaces, it would be very informative to add a figure or table comparing this robot with others reported in the literature in terms of size, turning radius, degrees of freedom, locomotion performance, and power consumption.
(9)The demonstration of multi-robot position-data sharing is impressive. Showing collaboration among three or more robots would make the scenario even more compelling.
(10)Some figures miss data labels, such as Fig. 2c and Fig. 2h; these should be clearly marked.

Reviewer #2

(Remarks to the Author)
This paper presents the design and experimental validation of a miniature legged robot driven by piezoelectric actuators. A key feature of the robot is its ability to move omnidirectionally without turning, achieved by selectively actuating one or two of its legs. The system is interesting and demonstrates a high level of integration among onboard sensors, actuators, and power components. The authors also provide several compelling demonstrations. Overall, the paper is well written and easy to follow. However, the reviewer has several comments that may help strengthen the paper:

- The authors emphasize the robot's ability to move without turning. However, it is unclear why minimizing the turning radius is critical for such a small robot. If the robot is already miniature in size, having a small turning radius might not pose a significant limitation. Please elaborate on the practical importance of this design goal.
- While the robot is claimed to achieve both micro- and macro-scale motion, the experimental videos show that its motion is not very straight. The trajectory appears to deviate, suggesting potential limitations in movement accuracy. Can the robot reliably follow a straight line? Please comment on the control or mechanical factors that contribute to this behavior.
- Since the robot's motion is vibration-based, it appears to require a smooth surface for locomotion. How would it perform on textured or uneven surfaces? Even small obstacles might prevent movement. Please discuss the robot's limitations and potential strategies for improving robustness on non-smooth terrains.

Minor Comments
• Figure 1(c-ii): The two legs are difficult to distinguish. Consider improving contrast or adjusting the view to make the leg structures clearer.
• Figure 2(b): The meaning of the measurements is unclear. Please label the reference frame (O–XYZ) relative to the leg to aid interpretation. It would also help to visually indicate the vibration pattern, rather than showing only a shaded region.
• Figure 2(c): Clarify that the three curves are projections onto the plane. Adding simple annotations and a schematic of the leg would make this figure more intuitive.
• Figure 3(a): The subfigures are currently difficult to interpret. The caption mentions "power," but only the top-left subfigure appears to show power data. Please clarify the meaning of the other subfigures and label the vertical axes (especially for the top-right and bottom-right plots).
• Application Justification: In the section discussing the robot's use under a microscope for scratch observation, it is unclear why a robot is needed if the microscope can directly observe the scratch. Please clarify the specific advantage or necessity of using the robot in this context.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I thank the authors for the revised manuscript and my concerns have been addressed in full. I recommend publication of this work.

Reviewer #2

(Remarks to the Author)
The reviewer appreciates the authors' substantial efforts in addressing the previous comments. All concerns have been satisfactorily resolved, and the reviewer has no further suggestions for improvement.

Responses to reviewers' comments (Reviewer: #1)

Comments to the Author:

This manuscript developed a novel miniature tripodal piezoelectric robot that demonstrates omnidirectional locomotion with a zero-radius turning capability. Notably, the design enables smooth transitions across a wide velocity range—from micrometer to centimeter scales—showcasing agile navigation in macroscopic environments and precise motion control in microscopic scenarios. The robot's circular symmetric architecture further allows compact, stackable integration of multiple functional modules. The authors also present closed-loop trajectory control and multi-robot cooperation capabilities. Overall, this work represents an advancement in structural design, steering performance, and system scalability of piezoelectric miniature robots. The idea and concept of the paper are excellent and novel, but the available data on this particular aspect appears to be insufficient. The authors are encouraged to consider the following suggestions for enhancing the manuscript in preparation for future submissions.

Response: We sincerely appreciate your professional comments and your recognition of our robot's novel design. We found your constructive suggestions regarding the sufficiency of the data to be highly valuable. In response to your suggestion, we have performed additional experiments and provided more comprehensive data to further validate the robot's performance. The manuscript has been carefully revised to incorporate these new results. Detailed responses to each of your comments are listed below, and all changes have been highlighted in yellow in the revised manuscript and Supplementary Information.

Comment 1. The authors claim that the robot exhibits exciting zero-radius omnidirectional turning capability. However, the experiments only demonstrate several specific turning angles. It is necessary to provide a complete set of turning performance data, covering rotation angles from 0° to 360° (with a suitable step size, e.g., 5°), to substantiate the claim.

Response: Thank you for your professional suggestions.

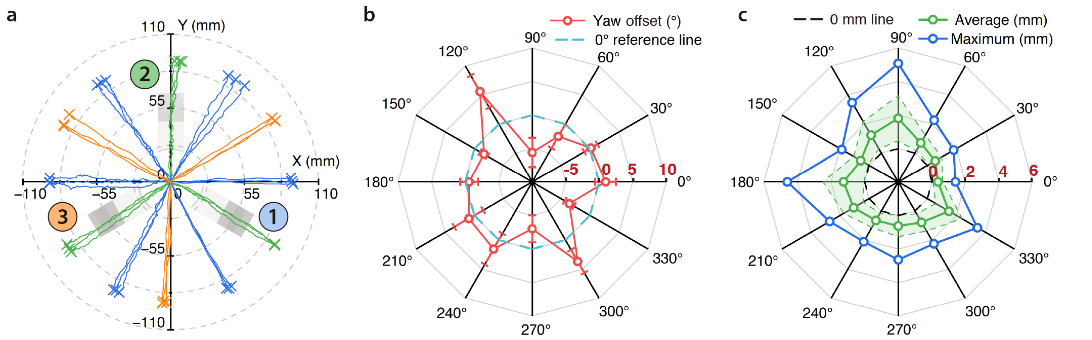
The robot possesses the capability for omnidirectional planar locomotion without the need for attitude adjustment. It is worth noting that the robot's motion does not involve "rotation," and consequently, there is no definition of "rotation angles" within its motion framework. The robot omnidirectional motion strategy includes two key approaches. (i) Tripodal radial actuation strategy. (ii) Trajectory interpolation algorithm. By combining these two aspects, the robot achieves omnidirectional planar locomotion capability without the need for attitude adjustment. The tripodal radial actuation strategy achieves six discrete motion directions symmetrically distributed along the circumference. Other non-conventional angles are realized through planar positioning, calculated via the trajectory interpolation algorithm.

In our original experimental tests, we validated effective locomotion along three fundamental and three synthetic directions, leveraging the single and dual actuation modes within the tripodal radial actuation strategy. Moreover, by modulating the driving force ratios of the two active legs in the optimal vibration state (OVS), we successfully demonstrated the synthesis of six new intermediate angles, achieving asymmetric locomotion directions. We conducted additional tests to evaluate the robot's omnidirectional synthetic motion capabilities across 12 directions with 30° increments (see **Supplementary Movie S3**, Part III). The corresponding characterization results are presented in **Supplementary Fig. 10**. For each direction, a 90-mm trajectory segment is extracted for analysis (**Supplementary Fig. 10a**). Orientation stability is evaluated via yaw deviation, defined as the angular difference between the theoretical target and the actual trajectory angle (ranging from 0° to 360° relative to the positive x-axis). The results show a minimum yaw offset of 0.2° during motion toward the 30° azimuth, while a maximum offset of 5.6° occurred at the 120° azimuth (**Supplementary Fig. 10b**). Furthermore, the trajectory deviation, including the average and maximum trajectory deviations, are employed to quantify trajectory-keeping capabilities (definitions provided in **Supplementary Fig. 10c** caption). The average deviation ranged from a minimum of 0.4 mm (0.4% relative to travel distance) at 0° to a maximum of 1.8 mm (2.0%) at 90°, while the maximum deviation spanned from 1.4 mm (1.6%) at 0° to 5.1 mm (5.7%) at 90°. These results demonstrate that the robot, governed by the tripodal radial actuation strategy, effectively achieves planar motion across 12 discretized directions (30° resolution) with trajectory-keeping capabilities.

Modification: We have added relevant descriptions and analyses to the manuscript.

*"By modulating the driving force ratios of the two active legs in OVS, we find that the tripodal radial actuation strategy can also synthesize six additional asymmetric motion angles. We characterize the robot's omnidirectional motion across 12 directions in 30° increments, including six asymmetric synthetic angles (see Part III in **Supplementary Movie S3**). For each direction, a 90-mm trajectory segment is extracted for analysis (**Supplementary Fig. 10a**). Orientation stability is evaluated via yaw offset, defined as the angular difference between the theoretical target and the actual trajectory angle (ranging from 0° to 360° relative to the positive x-axis). The results show a minimum yaw offset of 0.2° during motion toward the 30° azimuth, while a maximum offset of 5.6° occurred at the 120° azimuth (**Supplementary Fig. 10b**). Furthermore, the trajectory deviation, including the average and maximum trajectory deviations, are employed to quantify trajectory-keeping capabilities (definitions provided in **Supplementary Fig. 10c** caption). The average deviation ranges from a minimum of 0.4 mm (0.4% relative to travel distance) at 0° to a maximum of 1.8 mm (2.0%) at 90°, while the maximum deviation spans from 1.4 mm (1.6%) at 0° to 5.1 mm (5.7%) at 90°. These results demonstrate that the robot, governed by the tripodal radial actuation strategy, effectively achieves planar motion across 12 discretized directions (30° resolution) with trajectory-keeping capabilities."*

(please see page 11, line 260 of the revised manuscript, marked in yellow)



*"**Supplementary Fig. 10.***

Demonstration and evaluation of planar motion across twelve discrete directions. (a)** Trajectory map of the robot across 12 directions with 30° increments (each movement spans 90 mm; $n=3$ trials per direction). **(b)** Yaw offset across different motion directions. The yaw angle is defined as the orientation of the motion trajectory relative to the positive x-axis (ranging from 0° to 360°), and the offset represents the angular difference between the actual trajectory and the theoretical target angle. Minimum value: offset 0.2° at 30°, maximum value: offset 5.6° at 120°. **(c)** Evaluation of trajectory deviations across* *motion directions. For a 90-mm trajectory segment, the average deviation ranges from a minimum of 0.4 mm (0.4% relative to travel distance) at 0° to a maximum of 1.8 mm (2.0%) at 90°. The maximum deviation ranges from a minimum of 1.4 mm (1.6%) at 0° to a maximum of 5.1 mm (5.7%) at 90°. The green dashed area represents the standard deviation (SD).* ***Note: *Data obtained from a representative prototype (Prototype #2). The green dashed area represents the standard deviation (SD).* ***Note: Deviation metrics.*** *Trajectory precision is quantified by the average deviation ($D_{avg} = \frac{1}{n}\sum|d_i|$) and maximum deviation ($D_{max} = \max|d_i|$). Here, d_i represents the perpendicular distance of each actual trajectory point from the straight line connecting the segment's start and end points."*

(please see page 27 of the Supplementary Materials, Supplementary Fig. 10, marked in yellow)

Comment 2. The environmental adaptability of the miniature robot is also of great concern. Could the authors evaluate its performance in more challenging terrains, such as uneven surfaces, slopes, or gravel grounds? Furthermore, how well does the robot maintain its locomotion capability after compression or drop impacts?

Response: Thank you for your professional advices.

Beyond the glass substrate, the robot also demonstrates effective locomotion on marble and stainless-steel surfaces (see Part I in **Supplementary Movie S4**). The velocity characteristics across the six primary motion directions on these diverse contact interfaces are presented in **Supplementary Fig. 11**. The robot achieves a peak velocity of 30.66 mm/s on the glass substrate (actuation: #1 and #3; low-frequency bounce state (LBS): #2). In comparison, the maximum locomotion speeds on marble and stainless-steel surfaces reach 26.21 mm/s and 26.06 mm/s, respectively (actuation: #1 and #2; LBS: #3 for both cases). Moreover, the robot demonstrates steady ascent at a maximum incline angle of 1.5° (see Part III in **Supplementary Movie S4**). As the tilt angle increased to 2°, the robot exhibits stalling and failed to maintain movement on the incline.

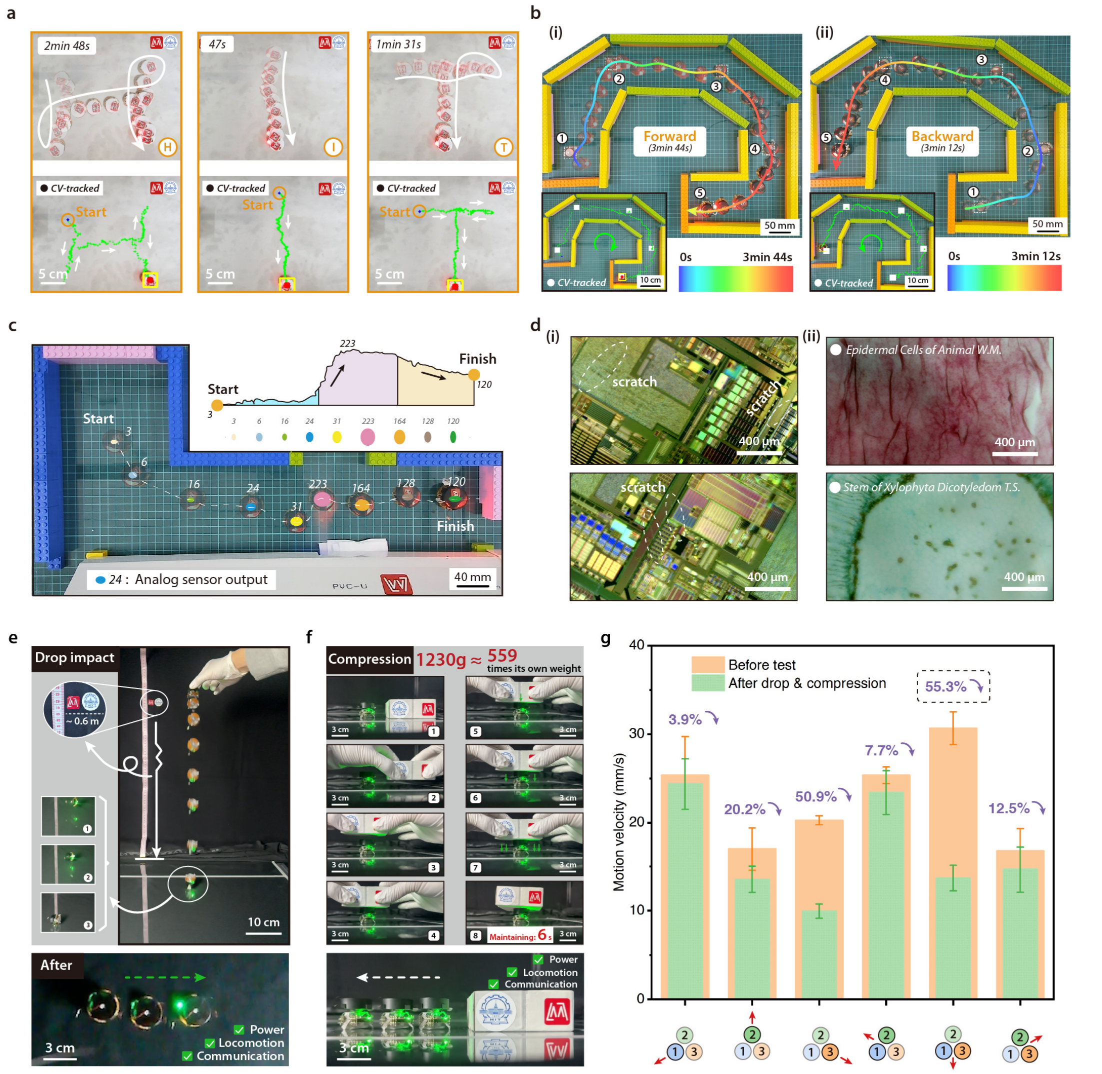
In fact, the vibration amplitudes of the robot's foot tips are typically on the micrometer scale (OY: 11.15 μm, OZ: 8.83 μm; see **Fig. 2b** in the main text). Such minute displacements are insufficient to support effective locomotion on uneven surfaces, such as gravel terrain. The tripodal robot uses 1st-order bending vibration as the operating mode, exciting a high-frequency in-plane oblique trajectory at the driving foot. The robot accomplishes microscopic movement by utilizing microscopic vibration on a plane, and high-speed macroscopic movement can be achieved through the accumulation of high-frequency microscopic small steps. This is a common characteristic of vibration-driven piezoelectric robots. Our work is situated within the context of precision robotics for applications in fields like micro-operation, where the operational environments (such as glass, marble, or stainless-steel surface) are flat. Nevertheless, we agree that considering the robot's ability to handle rougher surfaces is very valuable for expanding the application scenarios of piezoelectric robots. We have added the limitation concerning limited locomotion substrates and incorporated corresponding future work into the Discussion section to improve the comprehensiveness of the study.

Besides, we have performed additional robustness tests, including free-fall impact and compression tests. As shown in **Fig. 4e**, the robot is elevated to a height of 0.6 m and released in its normal working orientation for a free-fall impact onto the working surface. Despite flipping and rolling upon impact, the robot maintained normal power supply, communication, and locomotion functionality (see Part I in **Supplementary Movie S10**). Subsequently, a massive iron block, weighing approximately 559 times the original tripodal robot (2.2 g), is placed on the robot for 6 seconds (**Fig. 4f**). Although structural deformation occurred under the load, the robot partially recovers its original shape after the load is removed and retain its capability for untethered locomotion (see Part II in **Supplementary Movie S10**). The velocity characteristics across six primary directions are evaluated (**Fig. 4g**). While the robot preserves its locomotion capability in all directions, varying degrees of speed reduction are observed. The maximum reduction occurs in the synthetic direction of driving legs #1 and #3 (dual actuation mode), representing a decrease of approximately 55.3%. This is followed by the single actuation mode (actuation: #3; LBS: #1 and #2), which shows a 50.9% decline. These results suggest that driving leg #3 sustains relatively more severe damage during the two tests, which alters the overall vibration characteristics.

Modification: The experimental results and limitations have been added to the revised manuscript.
*"Beyond the glass substrate, the robot also demonstrates effective locomotion on marble and stainless-steel surfaces (see Part I in **Supplementary Movie S4**). The velocity characteristics across the six primary motion directions on these diverse contact interfaces are presented in **Supplementary Fig. 11**. The robot achieves a peak velocity of 30.66 mm/s on the glass substrate (actuation: #1 and #3; LBS: #2). In comparison, the maximum locomotion speeds on marble and stainless-steel surfaces reach 26.21 mm/s and 26.06 mm/s, respectively (actuation: #1 and #2; LBS: #3 for both cases)."*
(please see page 12, line 277 of the revised manuscript, marked in yellow)

*"Moreover, the robot demonstrates steady ascent at a maximum incline angle of 1.5° (see Part III in **Supplementary Movie S4**). As the tilt angle increased to 2°, the robot exhibits stalling and failed to maintain movement on the incline."*
(please see page 13, line 299 of the revised manuscript, marked in yellow)

*"To evaluate the robustness required for practical applications, the robot is elevated to a height of 0.6 m and released in its working posture for a free-fall impact onto the working surface (**Fig. 4e**). Despite flipping and rolling upon impact, the robot maintained normal power supply, communication, and locomotion functionality (see Part I in **Supplementary Movie S10**). Subsequently, a massive iron block, weighing approximately 559 times (1230 g) the original tripodal robot (2.2 g), is placed on the robot for 6 seconds (**Fig. 4f**). Although structural deformation occurred under the load, the robot partially recovers its original shape after the load is removed and retain its capability for untethered locomotion (see Part II in **Supplementary Movie S10**). Following the two sets of experiments, the degradation of velocity characteristics across the six primary directions is evaluated (**Fig. 4g**). While the robot preserves its locomotion capability in all directions, varying degrees of speed reduction are observed. The maximum reduction occurs in the synthetic direction of driving legs #1 and #3 (dual actuation mode), representing a decrease of approximately 55.3%. This is followed by the single actuation mode (actuation: #3; LBS: #1 and #2), which shows a 50.9% decline. These results suggest that driving leg #3 sustains relatively more severe damage during the two tests, which alters the overall vibration characteristics."*
(please see page 16, line 381 of the revised manuscript, marked in yellow)

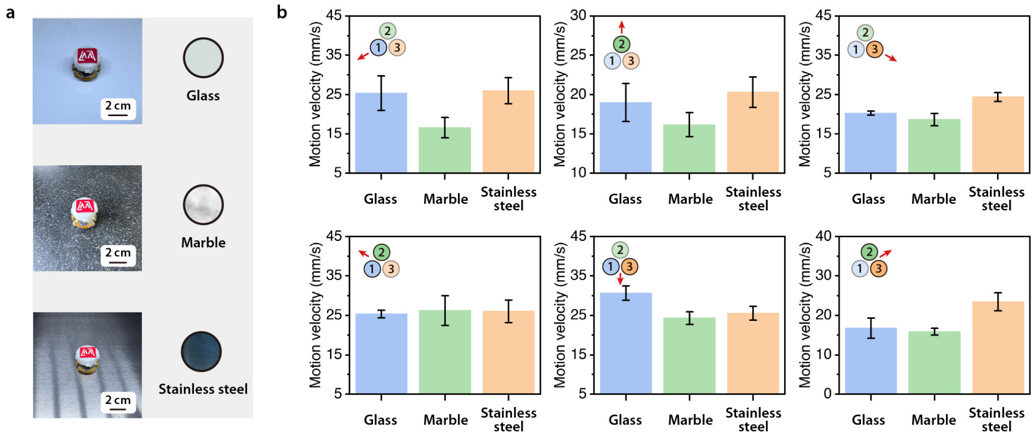


*“**Fig 4. Performance and applications for the robot. (a)** The untethered robot draws the trajectory of “H”, “I”, and “T” letters on the iron substrate. **(b)** The robot demonstrates bidirectional maze navigation and reports five magnetic checkpoints it passed using the onboard hall sensor. (i) Forward. (ii) Backward. **(c)** The gas leak inspection application. (Note: the values only represent the sensor analog values and are not actual gas concentrations) **(d)** The robot is capable of performing tasks in microscopic observation. (i) Wafer detection. (ii) Cell specimen observation (Epidermal Cells of Animal W.M. and Stem of Xylophyta Dicotyledon T.S.). **(e)** The robot (Prototype #2) remains functional after a free-fall impact from a height of 0.6 m onto a desktop. **(f)** The robot (Prototype #2) maintains normal locomotion after sustaining a compressive load of 1230 g (approximately 559 times its body weight). **(g)** Evaluation of velocity reduction after a drop impact and compressive loading (Prototype #2, 60 $V_{p\text{-}p}$; maximum reduction: 55.3% reduction; Actuation: #1 and #3, LBS: #2).”*

(please see page 37, line 777 of the revised manuscript, Fig. 4, marked in yellow)

“iii) Since the robot is driven by micro-amplitude vibrations of its foot tips, its operation has primarily been demonstrated on smooth, planar surfaces. Exploring and expanding the robot's locomotion capabilities on non-uniform surfaces is essential for broadening its potential application scenarios.”

(please see page 22, line 522 of the revised manuscript, marked in yellow)



*"**Supplementary Fig. 11.***

Performance of the robot on diverse contact substrates. (a) *Photographs of the robot situated on glass, marble, and stainless-steel surfaces.* ***(b)*** *The velocity characteristics across six motion directions on the three respective substrates. The robot achieves a peak velocity of 30.66 mm/s on the glass substrate (actuation: #1 and #3; LBS: #2). In comparison, the maximum locomotion speeds on marble and stainless-steel surfaces reach 26.21 mm/s and 26.06 mm/s, respectively (actuation: #1 and #2; LBS: #3 for both cases). Note: Data obtained from a representative prototype (Prototype #2, 60 $V_{p\text{-}p}$)."*

(please see page 28 of the Supplementary Materials, Supplementary Fig. 11, marked in yellow)

Comment 3. The demonstrated wide velocity range—from 3 cm/s to 5.63 μm/s—is impressive, showing potential for both macro- and micro-scale explorations. However, I am particularly concerned about the precision of velocity control, which directly determines the smoothness of task execution.

Response: We appreciate your positive assessment of our velocity range and the constructive comment regarding velocity control precision. Regarding velocity control precision, it is essential to distinguish between open-loop and closed-loop aspects. While the robot possesses closed-loop capabilities, the control architecture detailed in this study prioritizes the regulation of position rather than velocity. In our closed-loop experiments, the driving parameters (voltage and frequency) are kept constant, resulting in a fixed speed, while the control is achieved by switching the driving states (OVS or LBS) of the legs. Consequently, the discussion on velocity precision in this response focuses on the open-loop motion process.

This open-loop motion spans two distinct operational modes: macroscopic rapid planar motion and microscopic high-precision stepping.

- For the macroscopic rapid motion mode, velocity regulation is primarily achieved by adjusting the frequency and voltage of the driving signals. Since the robot utilizes piezoelectric actuators operating in resonance mode, there exists an optimal resonant frequency. Consequently, deviating from this resonant frequency (either by increasing or decreasing it) reduces the vibration energy of the driving legs, thereby indirectly lowering the robot's motion speed. Similarly, modulating the amplitude of the driving voltage can also be used to indirectly adjust the robot's velocity.
- For microscopic low-speed motion, intermittent excitation signals are generated by applying a signal mask to the original high-frequency excitation signal (~30 kHz), enabling the robot to perform stepping motion. The step size is regulated by adjusting the frequency (fixed at 10 Hz in this study) and the duty cycle (e.g., 1%, 2%) of the signal mask.

We employ a sensitivity metric to evaluate the velocity control resolution of these two motion modes. This metric is defined as the change in velocity resulting from a unit variation in the control parameter (e.g., 1 Hz, 1 V_{p-p}, or 1% duty cycle).

- For macroscopic rapid movements

As shown in **Fig. 3c**, we calculated the sensitivity of the robot's velocity to frequency ($mm \cdot s^{-1} \cdot Hz^{-1}$). In the three basic motion directions, the maximum frequency sensitivity is 0.52 $mm \cdot s^{-1} \cdot Hz^{-1}$. Since the experimental frequency interval is 10 Hz, the maximum amplitude of velocity variation is 5.2 mm/s. For synthetic motion directions, the robot operates in dual actuation mode, requiring two driving legs to be simultaneously in OVS. Regulating velocity via frequency is complex in this configuration. Therefore, we typically adjust the driving voltage to control speed. According to the voltage-velocity characteristic curve (**Fig. 3d** in the main text), the voltage sensitivities for the three synthetic directions are 0.26, 0.19, and 0.52 $mm \cdot s^{-1} \cdot V_{p-p}^{-1}$, respectively. With a minimum adjustment increment of 30 V_{p-p}, the corresponding maximum velocity variations are 7.8 mm/s, 5.7 mm/s, and 15.6 mm/s, respectively.

It is important to note that the robot requires a minimum excitation voltage of 30 V_{p-p} to initiate effective motion. This threshold results in minimum velocities of 7.73 mm/s (fundamental directions) and 12.74 mm/s (synthetic directions), indicating a dead zone in the macroscopic velocity. Therefore, the effective velocity regulation ranges are limited to 7.73 ~ 24.23 mm/s for the basic directions and 12.74 ~ 30.08 mm/s for the synthetic directions. We have added these details to the revised text to clarify the robot's motion characteristics.

- Regarding microscopic low-speed motion

As previously described, the robot's microscopic motion relies on the onboard circuit unit, which generates a signal mask based on the original continuous excitation signal to achieve intermittent step motion. To facilitate testing, we fixed the signal mask frequency at 10 Hz and varied only the duty cycle to modulate the robot's step size. We tested the relationship between step size and duty cycle across all six directions (refer to **Fig. 3e** in the main text) and calculated the duty cycle sensitivity, as shown in **Supplementary Fig. 14b**. The results indicate that the maximum average step size variation per 1% duty cycle increment is 1.38 μm, with an actual sensitivity range of 0.29 ~ 3.53 μm/1% duty cycle. This demonstrates that the control precision of the microscopic motion is at the micrometer level, exhibiting excellent controllability.

It should be clarified that while the robot's two motion modes (macroscopic rapid motion and microscopic stepping motion) are independent, this does not imply a complex switching logic. On the contrary, since the driving conditions for these two modes differ only in the type of driving signal, the switching cost is relatively low, ensuring operational convenience.

Additionally, we acknowledge that our previous description regarding "microscopic motion velocity" may have been potentially misleading. In the microscopic stepping mode, the robot is driven by intermittent high-frequency square wave signals. The motion process can be viewed as a sequence of "move-stop-move-stop" cycles. Strictly speaking, this is not continuous motion. Furthermore, the term "minimum motion velocity" is technically inaccurate, as the theoretical minimum velocity is zero when the robot is stationary. Consequently, we have revised the terminology in the manuscript: "microscopic low-speed motion" has been updated to "microscopic stepping motion", and "minimum motion velocity" has been corrected to "minimum step size".

We sincerely appreciate your insightful suggestion regarding velocity control precision. We recognize that high precision in velocity control is crucial for enhancing the robot's potential in practical applications. In our future work, we plan to explore velocity smoothing algorithms and implement closed-loop velocity control to further improve the controllability and the smoothness of task execution. We have added a discussion on these future directions in the **Discussion** section of the revised manuscript.

Modification: We have added the relevant description in the revised manuscript.

"To evaluate the rate of change in the excitation frequency-velocity relationship, we introduce a frequency sensitivity metric, defined as the magnitude of velocity variation per Hz. The results indicate that the sensitivity values for all three fundamental motion directions remain below 0.52 $mm{\cdot}s^{-1}{\cdot}Hz^{-1}$. Given that the minimum frequency adjustment increment used in the tests is 10 Hz, the achievable velocity control resolution is less than 5.2 mm/s."

(please see page 10, line 238 of the revised manuscript, marked in yellow)

"Notably, in the dual actuation mode (with two legs in OVS), controlling velocity through frequency modulation proves difficult. Consequently, we employ voltage adjustment to regulate the robot's motion speed."

(please see page 10, line 243 of the revised manuscript, marked in yellow)

"According to the voltage-velocity relationship, the voltage sensitivities for the three synthetic directions are 0.26, 0.19, and 0.52 $mm{\cdot}s^{-1}{\cdot}V_{p\text{-}p}^{-1}$. With a minimum adjustment increment of 30 $V_{p\text{-}p}$, the corresponding maximum velocity variations are 7.8 mm/s, 5.7 mm/s, and 15.6 mm/s, respectively."

(please see page 11, line 249 of the revised manuscript, marked in yellow)

"We observe that the minimum excitation voltage required for effective motion is 30 $V_{p\text{-}p}$. This corresponds to minimum velocities of 7.73 mm/s and 12.74 mm/s for the three fundamental directions and the three synthetic directions, respectively. Therefore, the velocity regulation ranges are 7.73 ~ 24.23 mm/s for the fundamental directions and 12.74 ~ 30.08 mm/s for the synthetic directions."

(please see page 11, line 255 of the revised manuscript, marked in yellow)

*"We utilize the derivative of the duty cycle-step size curve to define the step sensitivity (in μm/1% duty cycle), which serves to quantify the step control precision during micro-scale low-speed motion, as shown in **Supplementary Fig. 14b**. The maximum average step size variation per 1% duty cycle increment is 1.38 μm, with an actual sensitivity range of 0.29 ~ 3.53 μm/1% duty cycle (actuation: #3, LBS: #1 and #2)."*

(please see page 13, line 314 of the revised manuscript, marked in yellow)

"As the two modes differ only in the form of the driving signal, the switching overhead remains low, ensuring operational convenience."

(please see page 14, line 325 of the revised manuscript, marked in yellow)

"By switching between high-frequency continuous and intermittent excitation modes, the robot achieves a wide dynamic range, spanning from a macroscopic speed of 3 cm/s down to a microscopic step resolution of 0.56 μm, thereby enabling both rapid maze navigation and precise microscopic observation."

(please see page 1, line 22 of the revised manuscript, marked in yellow)

"ii) Cross-scale motion. By switching between high-frequency continuous and intermittent excitation modes, the robot achieves a wide dynamic range, spanning from a macroscopic speed of 3 cm/s down to a microscopic step resolution of 0.56 μm, thereby enabling both rapid maze navigation and precise microscopic observation. (such as wafer detection and cell specimen observation)."

(please see page 4, line 83 of the revised manuscript, marked in yellow)

*"**Figure. 3f**(i) presents a representative displacement curve of the stepping motion. Notably, the robot exhibits a stepping motion pattern, with its motion frequency matching the mask frequency (10 Hz, 10 steps per second). The robot achieves micrometer-level minimum step sizes in all six directions, with the minimum step size reaching 0.56 μm (Actuation: #1 and #2; LBS: #3), as shown in **Fig. 3f**(ii)."*

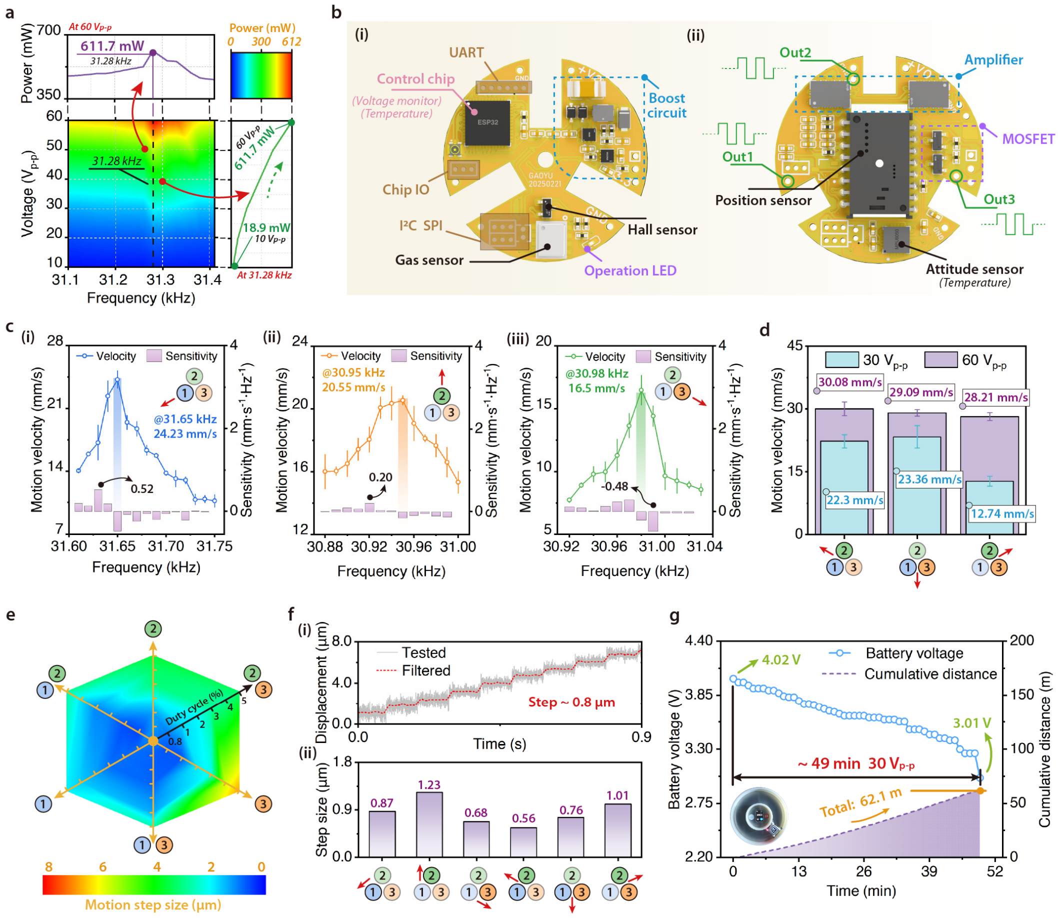
(please see page 13, line 319 of the revised manuscript, marked in yellow)

"Moreover, the robot can switch between a continuous mode with speeds up to 3 cm/s for maze navigation, and an intermittent mode with a minimum step size of 0.56 μm for precision microscopic observation. This cross-scale locomotion capability enables the robot to efficiently perform wide-range, multi-point chip defect inspections."

(please see page 20, line 486 of the revised manuscript, marked in yellow)

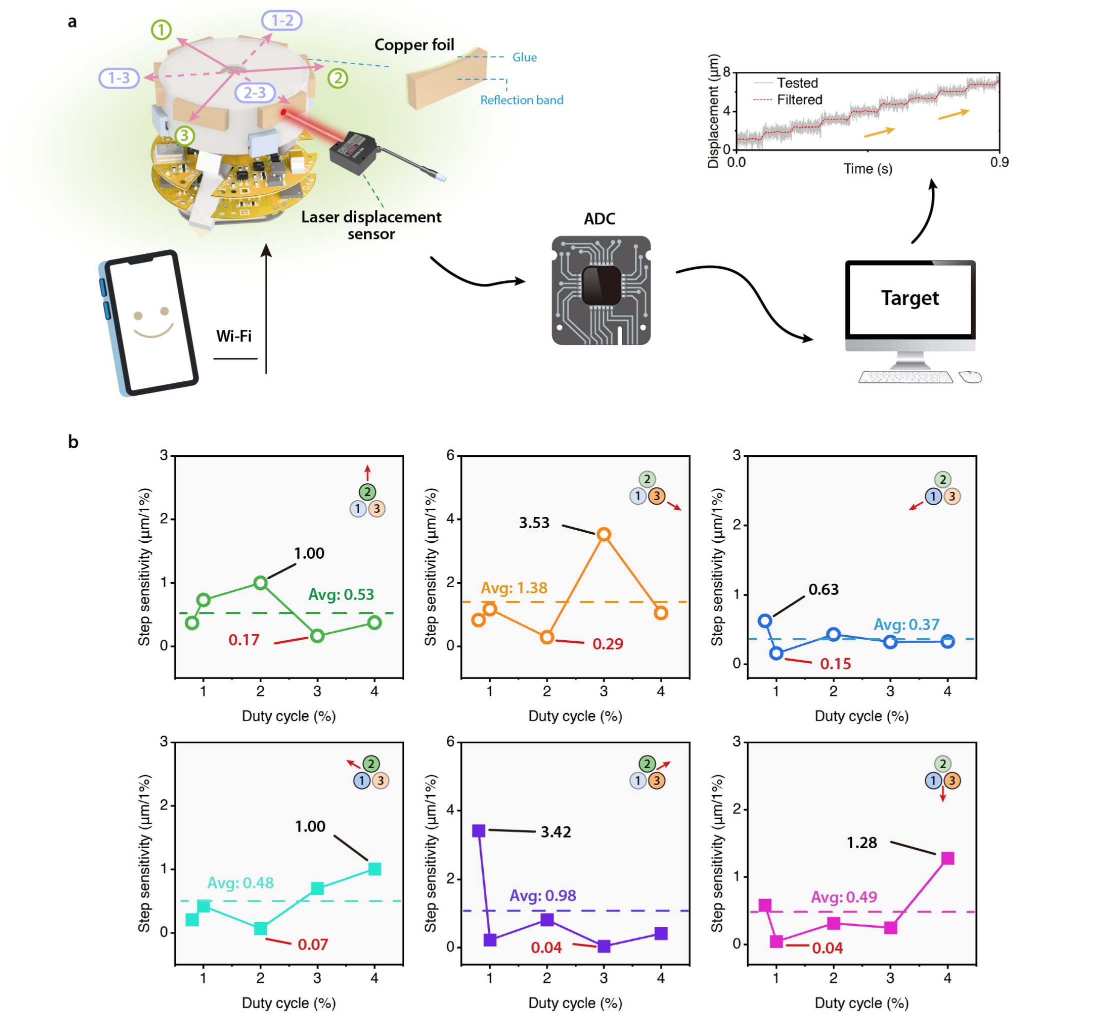
"iv) Developing velocity smoothing algorithms and implementing closed-loop velocity control systems to further enhance the robot's controllability and ensure the smoothness of task execution."

(please see page 22, line 534 of the revised manuscript, marked in yellow)



*"**Fig 3. Integration design and locomotion characteristics.** **(a)** Power consumptions of the tethered robot (single actuation unit, maximum value: 31.28 kHz, 611.7 mW, 60 $V_{p\text{-}p}$). Note: the panels show the frequency response at 60 $V_{p\text{-}p}$ (top left), the voltage response at 31.28 kHz (bottom right), and the overall power contour map (bottom left) with its corresponding color scale (top right).* ***(b)** Layout and design of the onboard circuit unit. (i) Top board. (ii) Bottom board.* ***(c)** Velocity and velocity sensitivity versus frequency for the untethered robot in single actuation mode (30 $V_{p\text{-}p}$). (i) Actuation: #1, LBS: #2 and #3. (ii) Actuation: #2, LBS: #1 and #3. (iii) Actuation: #3, LBS: #1 and #2. (Error bars: three repeated motion trials.)* ***(d)** The relationship between velocity and voltage in dual actuation mode. (Error bars: three repeated motion trials.)* ***(e)** The relationship between motion step size and duty cycle in the six directions. (f) Minimum step size in stepping motion. (i) An example displacement data. (ii) Minimum microscopic motion step size in six directions (duty cycle: 0.8%, minimum step size: 0.56 μm).* ***(g)** The robot exhibits long endurance over 49 minutes, covering a cumulative travel distance of 62.1 m (initial voltage: 4.02 V, 180 mAh, 30 $V_{p\text{-}p}$)."*

(please see page 35, line 762 of the revised manuscript, Fig. 3, marked in yellow)



*"**Supplementary Fig. 14.***

***Microscopic step motion testing platform and step sensitivity evaluation. (a)** Schematic of the system. **(b)** Step sensitivity in six basic directions. Note: this metric is used to quantify the variation rate of the robot's step size, expressed as the amount of step-size change per 1% duty-cycle increment."*

(please see page 31 of the Supplementary Materials, Supplementary Fig. 14, marked in yellow)

Comment 4. From Fig. 3c and 3d, the velocity difference between the two-leg and one-leg motion modes appears marginal. Can I understand that the two-leg motion consumes approximately twice the power of the one-leg mode? If so, does this imply that the robot's energy efficiency decreases in certain motion directions?

Response: Thank you for this valuable comment. Although the dual actuation mode involves two active driving legs in OVS state compared to only one in the single actuation mode, the total power consumption does not exhibit a simple twofold increase (approximately 1.27 times). The power consumption of the tethered tripodal robot is characterized using a digital power meter (WT210, Yokogawa, Japan). **Supplementary Fig. 8a** illustrates the power consumption of a single driving leg as a function of excitation voltage in both OVS and LBS. At a constant frequency, the power consumption of the driving leg correlates positively with the applied voltage. Specifically, the maximum power consumption in LBS is measured at 313.8 mW (60 $V_{p\text{-}p}$). At an operating voltage of 30 $V_{p\text{-}p}$, the power consumption for the OVS and LBS is 154.3 mW and 74.0 mW, respectively.

- For the single actuation mode, the total power consumption is calculated as:

$$P_{\text{total}} = P_{\text{OVS}} + 2P_{\text{LBS}} \approx 302.3 \text{ mW} \quad (1)$$

- For the dual actuation mode, the total power consumption is calculated as:

$$P_{\text{total}} = 2P_{\text{OVS}} + P_{\text{LBS}} \approx 382.6 \text{ mW} \quad (2)$$

The power consumption in the dual actuation mode is approximately 1.27 times that of the single actuation mode.

Furthermore, our findings align with your hypothesis regarding the degradation of energy efficacy. Due to the nature of vector-summation-based motion, the robot inevitably experiences energy dissipation in certain resultant directions. To validate this hypothesis, we conducted a comprehensive analysis of the robot's motion efficacy. The resultant locomotion direction of the proposed tripodal robot is determined by the vector synthesis of the propulsive forces generated by the three driving legs (**Supplementary Fig. 8b**). The motion coordination is categorized into two modes:

➢ Single actuation mode (**Supplementary Fig. 8b(i)**):

Two driving legs are configured in LBS, while one leg serves as the primary actuator in OVS. Since the LBS is a friction-reduction mode, its driving force is assumed to be negligible (indicated by green dashed circles), whereas the leg in OVS provides the effective propulsive force (indicated by the blue arrow).

➢ Dual actuation mode (**Supplementary Fig. 8b(ii)**):

The robot's heading is determined by the resultant vector of two legs operating in OVS. Under the assumption of equal driving forces from both OVS legs, the motion direction bisects the angle between them. For locomotion in arbitrary directions, the heading is synthesized by modulating the ratio of the driving force magnitudes between the two legs in OVS.

As shown in **Supplementary Fig. 8c**, the robot's workspace is divided into six symmetrical sectors. The global coordinate system is defined by aligning the X-axis with the longitudinal axis of leg #1. The locomotion heading, defined by the yaw angle θ in [0°, 360°), is measured counter-clockwise from the positive X-axis. Taking the motion direction within the first sector (θ in [0°, 60°)) as an example (indicated by the red arrow in **Supplementary Fig. 8c**), the motion strategy is implemented as follows:

- Main leg: Leg #1 operates at full power as the primary propulsion source, with its normalized output force defined as $u_1 = 1$.

- Auxiliary leg: Leg #2 serves as an auxiliary actuator, providing a modulated driving force with a normalized coefficient $u_2 = K(\theta)$.
- Friction-reduction leg: Leg #3 is set to the LBS state, contributing negligible propulsive force $u_3 = 0$.

Based on the Law of Sines within the force vector triangle, the relationship between the required yaw angle θ and the driving coefficient $K(\theta)$ of the auxiliary leg can be derived. To synthesize a resultant motion at angle θ, the coefficient $K(\theta)$ is determined by:

$$K(\theta) = \frac{\sin(\theta)}{\sin(120° - \theta)}, \quad \theta \in [0°, 120°) \tag{3}$$

The modulation coefficient $K(\theta)$ is constrained within the range [0, 1]. This coefficient facilitates a seamless transition between the two previously defined actuation modes: when $\theta = 0°$, $K = 0$ (corresponding to single actuation mode), and when $\theta = 60°$, $K = 1$ (corresponding to dual actuation mode). Taking the 30 $V_{p\text{-}p}$ excitation voltage as a representative case, we assume that the effective propulsive force is linearly proportional to the electrical power consumed for mechanical work. Based on the experimental data, the power consumption model for a single driving leg can be formulated as follows:

$$P_{\text{leg}}(u) = \begin{cases} P_{\text{LBS}}, & u \approx 0 \\ P_{\text{LBS}} + u \cdot (P_{\text{OVS}} - P_{\text{LBS}}), & 0 < u \leq 1 \end{cases} \tag{4}$$

The power consumption of a single driving leg, denoted as $P_{\text{leg}}(u)$, is modeled as a function of the control coefficient u. For a constant excitation voltage of 30 $V_{p\text{-}p}$, let P_{LBS} represent the power consumption in LBS (74.0mW) and P_{OVS} represent the power consumption in OVS (154.3 mW). It can be inferred from the above model that the robot incurs a baseline power cost, defined by P_{LBS}, even when no effective propulsive force is generated. This baseline represents the essential energetic cost of maintaining the friction-reduction state. The total system power consumption, P_{total}, is calculated as the scalar summation of the power dissipated by the three independent driving legs:

$$P_{\text{total}}(\theta) = \sum_{i=1}^{3} P_{\text{leg}}(u_i) = P_{\text{OVS}} + 2P_{\text{LBS}} + (P_{\text{OVS}} - P_{\text{LBS}}) \cdot K(\theta) \tag{5}$$

While the preceding analysis focuses on the first sector, the methodology remains consistent across all six motion sectors by adjusting the range of the yaw angle θ and the assignment of the driving coefficient $K(\theta)$ to the respective auxiliary legs. **Supplementary Fig. 8d** provides a comprehensive summary of the target angles, driving coefficients, and the corresponding operational states of each leg across all sectors.

For force characterization, we normalize the resultant driving force based on the output of a single leg in the single actuation mode ($\theta = 0°$). Under this benchmark, only Leg #1 generates propulsive force. Consequently, the normalized magnitude of the synthesized resultant force, $\left|F_{syn}(\theta)\right|_{\text{norm}}$, is determined by the following vector summation:

$$\begin{aligned} \left|F_{\text{syn}}(\theta)\right|_{\text{norm}} &= \sqrt{1 + K(\theta)^2 + 2K(\theta)\cos(120°)} \\ &= \sqrt{1 - K(\theta) + K(\theta)^2} \end{aligned} \tag{6}$$

The normalized driving force coefficient across the full $0° \sim 360°$ range is shown in **Supplementary Fig. 8e**. The analysis reveals that the resultant driving force reaches its local maxima at the three fundamental actuation directions and the three primary symmetric synthesis directions. Conversely, the minimum driving force coefficient (approximately 0.87) occurs at the six offset synthesis angles (e.g., 30°).

Similarly, **Supplementary Fig. 8f** presents the total power consumption, normalized to the baseline power of the single actuation mode at $\theta = 0°$. Within each sector, the power consumption exhibits an approximately linear relationship with the yaw angle. The normalized power consumption coefficient reaches its peak of 1.27 at the dual-leg synthesis directions.

To quantitatively evaluate the trade-off between propulsive output and energetic investment across different headings, we define a dimensionless parameter, motion efficacy ($\eta(\theta)$), expressed as:

$$\eta(\theta) = \frac{\left|F_{\text{syn}}(\theta)\right|_{\text{norm}}}{P_{\text{total}}(\theta)_{\text{norm}}} \tag{7}$$

The motion efficacy of the robot across different headings is shown in **Supplementary Fig. 8g**. The results indicate that the robot achieves its peak locomotion efficacy at the three fundamental directions (i.e., 0°, 120°, and 240°), where only a single leg operates in OVS. In the dual actuation mode (e.g., 60°), the efficacy factor drops to 0.79 due to the increased power demand of concurrent leg activation. Furthermore, the efficacy reaches its minimum of approximately 0.75 at the offset synthesis angles, such as in the vicinity of 40° and 80°.This degradation in motion efficacy at synthesis angles can be attributed to the geometric projection loss inherent in vector summation. In the dual actuation mode and its adjacent offset angles, a portion of the generated driving force from each active leg is neutralized through internal cancellation along the non-target axes, whereas the total electrical power consumption scales cumulatively with the number of active actuators. Consequently, the rate of increase in power demand exceeds the growth of the resultant propulsive force, leading to a diminished efficacy profile compared to the single-leg actuation at fundamental angles.

It is worth noting that the motion efficacy model presented here, developed using the 30 $V_{p\text{-}p}$ case as a benchmark, incorporates several simplifying assumptions. Specifically, the model assumes a linear proportionality between the electrical power consumption and the resultant mechanical thrust, while neglecting potential non-linear effects such as interfacial friction heat dissipation, complex leg-ground interactions, and phase-dependent interference between multiple actuators.

Modification: We have added the relevant description in the manuscript and the supplementary materials. *"Additionally, the power consumption of a single driving leg in LBS is measured at 313.8 mW and 74.0 mW under excitation voltages of 60 $V_{p\text{-}p}$ and 30 $V_{p\text{-}p}$, respectively (**Supplementary Fig. 8a**). We analyze the motion efficacy of the robot across various directions and observe a decline in energy utilization at specific yaw angles (**Supplementary Note 5**). The reduced efficacy at synthesis angles results from vector cancellation combined with cumulative power consumption."*

(please see page 9, line 218 of the revised manuscript, marked in yellow)

*"**Supplementary Note 5***

Motion efficacy analysis

*The power consumption of the tethered tripodal robot is characterized using a digital power meter (WT210, Yokogawa, Japan). **Supplementary Fig. 8a** illustrates the power consumption of a single driving leg as a function of excitation voltage in both OVS and LBS. At a constant frequency, the power consumption of the driving leg correlates positively with the applied voltage. Specifically, the maximum power consumption in LBS is measured at 313.8 mW (60 $V_{p\text{-}p}$). At an operating voltage of 30 $V_{p\text{-}p}$, the* *power consumption for OVS and LBS is 154.3 mW and 74.0 mW, respectively. The resultant locomotion direction of the proposed tripodal robot is determined by the vector synthesis of the propulsive forces generated by the three driving legs (**Supplementary Fig. 8b**). The motion coordination is categorized into two modes:*

- *Single actuation mode (**Supplementary Fig. 8b(i)**):*

Two driving legs are configured in LBS, while one leg serves as the primary actuator in OVS. Since the LBS is a friction-reduction mode, its driving force is assumed to be negligible (indicated by green dashed circles), whereas the leg in OVS provides the effective propulsive force (indicated by the blue arrow).

- *Dual actuation mode (**Supplementary Fig. 8b(ii)**):*

The robot's heading is determined by the resultant vector of two legs operating in OVS. Under the assumption of equal driving forces from both OVS legs, the motion direction bisects the angle between them. For locomotion in arbitrary directions, the heading is synthesized by modulating the ratio of the driving force magnitudes between the two legs in OVS.

*As shown in **Supplementary Fig. 8c**, the robot's workspace is divided into six symmetrical sectors. The global coordinate system is defined by aligning the X-axis with the longitudinal axis of leg #1. The locomotion heading, defined by the yaw angle θ in [0°, 360°), is measured counter-clockwise from the positive X-axis. Taking the motion direction within the first sector (θ in [0°, 60°)) as an example (indicated by the red arrow in **Supplementary Fig. 8c**), the steering strategy is implemented as follows:*

- *Main leg: Leg #1 operates at full power as the primary propulsion source, with its normalized output force defined as $u_1 = 1$.*
- *Auxiliary leg: Leg #2 serves as an auxiliary actuator, providing a modulated driving force with a normalized coefficient $u_2 = K(\theta)$.*
- *Friction-reduction leg: Leg #3 is set to the LBS state, contributing negligible propulsive force $u_3 = 0$.*

Based on the Law of Sines within the force vector triangle, the relationship between the required yaw angle θ and the driving coefficient $K(\theta)$ of the auxiliary leg can be derived. To synthesize a resultant motion at angle θ, the coefficient $K(\theta)$ is determined by:

$$K(\theta) = \frac{\sin(\theta)}{\sin(120° - \theta)}, \quad \theta \in [0°, 120°) \tag{8}$$

The modulation coefficient $K(\theta)$ is constrained within the range [0, 1]. This coefficient facilitates a seamless transition between the two previously defined actuation modes: when $\theta = 0°$, $K = 0$ (corresponding to single actuation mode), and when $\theta = 60°$, $K = 1$ (corresponding to dual actuation mode). Taking the 30 $V_{p\text{-}p}$ excitation voltage as a representative case, we assume that the effective propulsive force is linearly proportional to the electrical power consumed for mechanical work. Based on the experimental data, the power consumption model for a single driving leg can be formulated as follows:

$$P_{\text{leg}}(u) = \begin{cases} P_{\text{LBS}}, & u \approx 0 \\ P_{\text{LBS}} + u \cdot (P_{\text{OVS}} - P_{\text{LBS}}), & 0 < u \le 1 \end{cases} \tag{9}$$

The power consumption of an individual driving leg, denoted as $P_{\text{leg}}(u)$, is modeled as a function of the control coefficient u. For a constant excitation voltage of 30 $V_{p\text{-}p}$, let P_{LBS} represent the power consumption in LBS (74.0mW) and P_{OVS} represent the power consumption in OVS (154.3 mW).

It can be inferred from the aforementioned model that the robot incurs a baseline power cost, defined by

P_{LBS}, even when no effective propulsive force is generated. This baseline represents the essential energetic cost of maintaining the friction-reduction state. The total system power consumption, P_{total}, is calculated as the scalar summation of the power dissipated by the three independent driving legs:

$$P_{\mathrm{total}}(\theta) = \sum_{i=1}^{3} P_{\mathrm{leg}}(u_i) = P_{\mathrm{OVS}} + 2P_{\mathrm{LBS}} + (P_{\mathrm{OVS}} - P_{\mathrm{LBS}}) \cdot K(\theta) \quad (10)$$

*While the preceding analysis focuses on the first sector, the methodology remains consistent across all six motion sectors by adjusting the range of the yaw angle θ and the assignment of the driving coefficient $K(\theta)$ to the respective auxiliary legs. **Supplementary Fig. 8d** provides a comprehensive summary of the target angles, driving coefficients, and the corresponding operational states of each leg across all sectors.*

To evaluate the energetic cost of different locomotion strategies at 30 $V_{p\text{-}p}$, we compare the two primary actuation modes:

- *For the single actuation mode ($\theta = 0°$, $K = 0$), the total power consumption is calculated as:*

$$P_{\mathrm{total}} = P_{\mathrm{OVS}} + 2P_{\mathrm{LBS}} \approx 302.3\ \mathrm{mW} \quad (11)$$

- *For the dual actuation mode ($\theta = 60°$, $K = 1$), the total power consumption is calculated as:*

$$P_{\mathrm{total}} = 2P_{\mathrm{OVS}} + P_{\mathrm{LBS}} \approx 382.6\ \mathrm{mW} \quad (12)$$

The power consumption in the dual actuation mode is approximately 1.27 times that of the single actuation mode. For force characterization, we normalize the resultant driving force based on the output of a single leg in the single actuation mode ($\theta = 0°$). Under this benchmark, only Leg #1 generates propulsive force. Consequently, the normalized magnitude of the synthesized resultant force, $\left|F_{syn}(\theta)\right|_{\mathrm{norm}}$, is determined by the following vector summation:

$$\begin{aligned} \left|F_{\mathrm{syn}}(\theta)\right|_{\mathrm{norm}} &= \sqrt{1 + K(\theta)^2 + 2K(\theta)\cos(120°)} \\ &= \sqrt{1 - K(\theta) + K(\theta)^2} \end{aligned} \quad (13)$$

*The normalized driving force coefficient across the full 0° ~ 360° range is shown in **Supplementary Fig. 8e**. The analysis reveals that the resultant driving force reaches its local maxima at the three fundamental actuation directions and the three primary symmetric synthesis directions. Conversely, the minimum driving force coefficient (approximately 0.87) occurs at the six offset synthesis angles (e.g., 30°). Similarly, **Supplementary Fig. 8f** presents the total power consumption, normalized to the baseline power of the single actuation mode at $\theta = 0°$. Within each sector, the power consumption exhibits an approximately linear relationship with the yaw angle. The normalized power consumption coefficient reaches its peak of 1.27 at the dual-leg synthesis directions.*

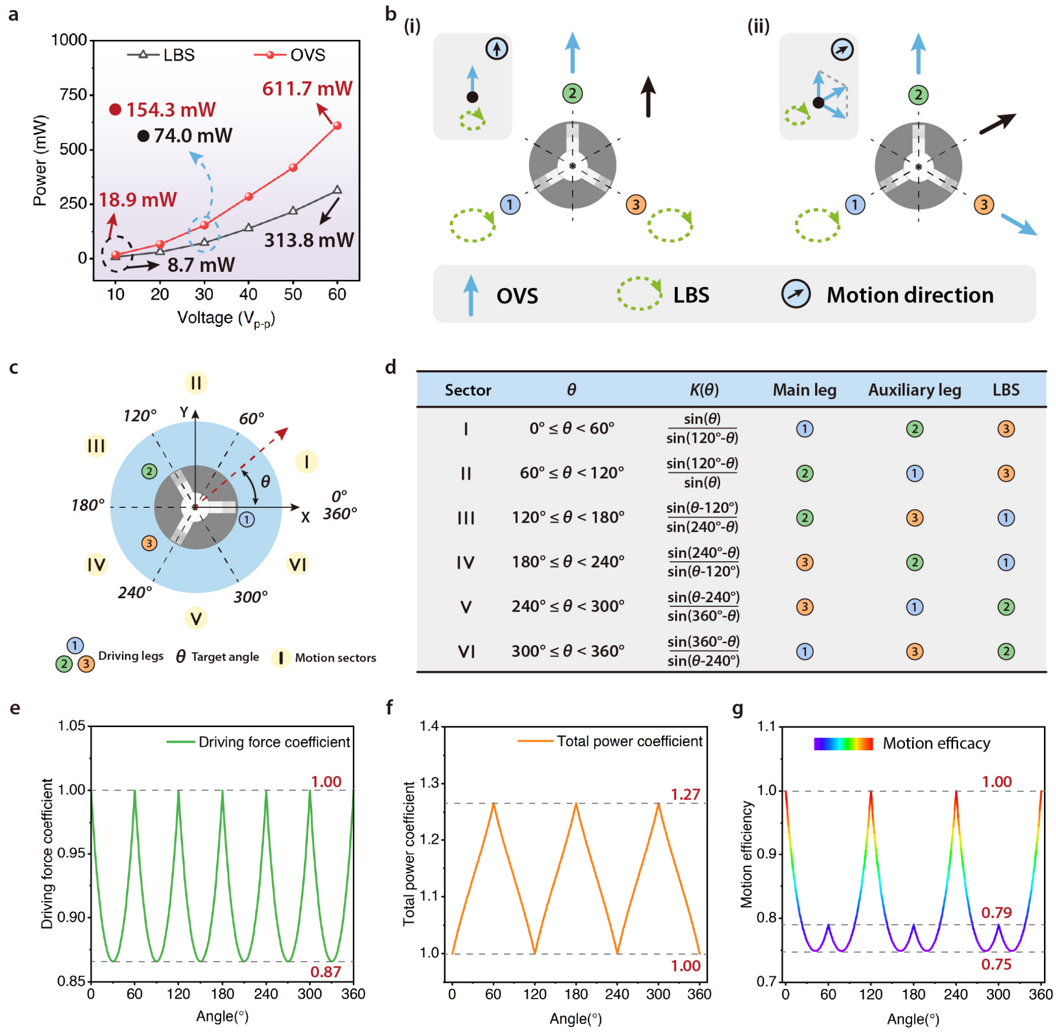
To quantitatively evaluate the trade-off between propulsive output and energetic investment across different headings, we define a dimensionless parameter, motion efficacy ($\eta(\theta)$), expressed as:

$$\eta(\theta) = \frac{\left|F_{\mathrm{syn}}(\theta)\right|_{\mathrm{norm}}}{P_{\mathrm{total}}(\theta)_{\mathrm{norm}}} \quad (14)$$

*The motion efficacy of the robot across different headings is illustrated in **Supplementary Fig. 6g**. The results indicate that the robot achieves its peak locomotion efficacy at the three fundamental directions (i.e., 0°, 120°, and 240°), where only a single leg operates in OVS. In the dual actuation mode (e.g., 60°), the efficacy factor drops to 0.79 due to the increased power demand of concurrent leg activation. Furthermore, the efficacy reaches its minimum of approximately 0.75 at the offset synthesis angles, such as in the vicinity of 40° and 80°. This degradation in motion efficacy at synthesis angles can be attributed* *to the geometric projection loss inherent in vector summation. In the dual actuation mode and its adjacent offset angles, a portion of the generated driving force from each active leg is neutralized through internal cancellation along the non-target axes, whereas the total electrical power consumption scales cumulatively with the number of active actuators. Consequently, the rate of increase in power demand exceeds the growth of the resultant propulsive force, leading to a diminished efficacy profile compared to the single-leg actuation at fundamental angles.*

It is worth noting that the motion efficacy model presented here, developed using the 30 $V_{p\text{-}p}$ case as a benchmark, incorporates several simplifying assumptions. Specifically, the model assumes a linear proportionality between the electrical power consumption and the resultant mechanical thrust, while neglecting potential non-linear effects such as interfacial friction heat dissipation, complex leg-ground interactions, and phase-dependent interference between multiple actuators."

(please see page 10 of the Supplementary Materials, Supplementary Note 5, marked in yellow)



Sector	θ	$K(\theta)$	Main leg	Auxiliary leg	LBS
I	$0° \le \theta < 60°$	$\frac{\sin(\theta)}{\sin(120°-\theta)}$	1	2	3
II	$60° \le \theta < 120°$	$\frac{\sin(120°-\theta)}{\sin(\theta)}$	2	1	3
III	$120° \le \theta < 180°$	$\frac{\sin(\theta-120°)}{\sin(240°-\theta)}$	2	3	1
IV	$180° \le \theta < 240°$	$\frac{\sin(240°-\theta)}{\sin(\theta-120°)}$	3	2	1
V	$240° \le \theta < 300°$	$\frac{\sin(\theta-240°)}{\sin(360°-\theta)}$	3	1	2
VI	$300° \le \theta < 360°$	$\frac{\sin(360°-\theta)}{\sin(\theta-240°)}$	1	3	2

"Supplementary Fig. 8.

Motion efficacy analysis of the robot. (a) *Power consumption of the driving leg as a function of voltage under different modes (OVS and LBS).* ***(b)*** *Schematic of the driving leg states and resultant force synthesis in two actuation modes.* ***(c)*** *Definition of the coordinate frames and sectors in the robot's motion efficacy model.* ***(d)*** *Summary of target angles (θ), driving coefficients ($K(\theta)$), and the corresponding states of different legs across various motion sectors.* ***(e)*** *Normalized driving force* *coefficient across the 0° ~ 360° range of motion directions (normalized to the driving force at a 0° motion angle in single actuation mode).* ***(f)*** *Normalized total power consumption coefficient across the 0° ~ 360° range of motion directions (normalized to the power consumption at a 0° motion angle in single actuation mode).* ***(g)*** *Motion efficacy of the robot across different motion directions from 0° ~ 360°."*

(please see page 25 of the Supplementary Materials, Supplementary Fig. 8, marked in yellow)

Comment 5. The section titled Multi-sensor Fusion Application may be slightly overstated. The experiment primarily demonstrates closed-loop trajectory tracking using a hand-drawn path, relying on an external camera and an onboard attitude sensor—only two sensors without redundancy. A more accurate title or clarification of the "fusion" scope is recommended.

Response: Thank you for this valuable comment, and we agree with your assessment regarding the terminology. In application demonstrations, the robot utilizes an external optical position sensor and an onboard attitude sensor to record real-time relative displacement and angular deviation (for attitude compensation), respectively. Our original intention was to highlight the system's miniaturization and high level of integration. We acknowledge that this differs from the definition of "sensor fusion," which typically implies algorithmic complementarity, mutual enhancement, and redundancy.

To describe the robot's design characteristics and the application experiment more accurately, we have changed the section title from "The multi-sensor fusion application demonstrations" to "Application of multi-sensor integration." Furthermore, we have replaced the term "fusion" with "integration" throughout the manuscript. We have also thoroughly checked and corrected relevant descriptions in both the main text and the Supplementary Materials. Moreover, we find your suggestion to be highly valuable and inspiring. We recognize that true multi-sensor data fusion is critical for enhancing the practical capabilities of robotic systems. Similarly, this has been identified as a key direction for future research.

Modification: We have revised the manuscript and the relevant sections in supplementary materials.

"... a long endurance over 49 minutes and applications of multi-sensor integration (including autonomous navigation and multi-robot coordination)."

(please see page 4, line 88 of the revised manuscript, marked in yellow)

*"**Application of multi-sensor integration**"*

(please see page 17, line 397 of the revised manuscript, marked in yellow)

"We propose a motion strategy based on multi-sensor integration, where complementary sensing capabilities are coordinated to automatically determine the required driving direction vector, enabling autonomous navigation and synchronized multi-robot operation."

(please see page 17, line 400 of the revised manuscript, marked in yellow)

"The robot achieves untethered motion and a long endurance (over 49 minutes), enabling applications based on multi-sensor integration."

(please see page 20, line 492 of the revised manuscript, marked in yellow)

*"**Fig 5. Demonstration of multi-sensor integration.**"*

(please see page 39, line 792 of the revised manuscript, Fig.5, marked in yellow)

Minor Comments

Comment 6. The robot's endurance of 20 minutes is reported, but it is difficult to assess whether this represents an advantage. It may be more convincing to report the total travel distance achievable under full battery conditions.

Response: We thank you for this valuable comment.

To evaluate the endurance and maximum range, four guide pulleys were installed around the robot, constraining its movement to a circular path with a radius of 65.5 mm. Under an excitation voltage of 30 $V_{p\text{-}p}$, the robot operated continuously for 49 minutes starting from a fully charged state (4.02 V) until it ceased motion at a terminal battery voltage of 3.01 V (**Fig. R1**, **Supplementary Movie S5**). During this period, the robot completed 151.1 laps, translating to a total travel distance of 62.1 m. This long endurance demonstrates the potential for multi-sensor integration and extended autonomous applications.

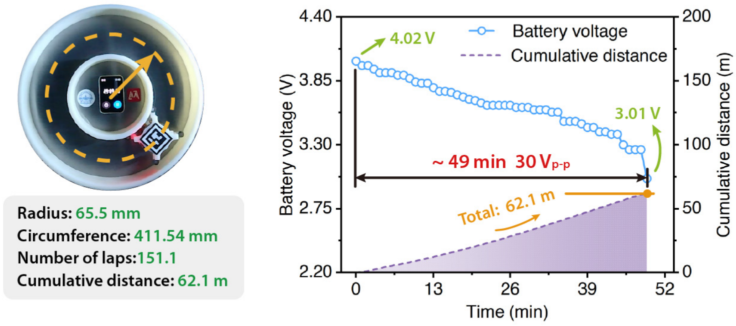


Fig R1. Robot endurance test results. The robot exhibits long endurance over 49 minutes, covering a cumulative travel distance of 62.1 m (initial voltage: 4.02 V, 180 mAh, 30 $V_{p\text{-}p}$).

Modification: The relevant content has been added to and updated in the revised manuscript. The content corresponding to **Fig R1** has been updated in **Fig. 3g**.

"iii) Miniaturization and high integration. Benefiting from the integration of actuation, communication, control, sensing (five types of sensors) and power supply units into the robot's body, the robot demonstrates untethered motion, a long endurance over 49 minutes and applications of multi-sensor integration (including autonomous navigation and multi-robot coordination)."

(please see page 4, line 88 of the revised manuscript, marked in yellow)

*"Moreover, powered by a 180 mAh lithium battery, the robot sustained wide-range macroscopic locomotion (30 $V_{p\text{-}p}$) for over 49 minutes, achieving a cumulative travel distance of 62.1 m (see **Supplementary Movie S5**)."*

(please see page 14, line 326 of the revised manuscript, marked in yellow)

"The robot achieves untethered motion and a long endurance (over 49 minutes), enabling applications based on multi-sensor integration."

(please see page 21, line 498 of the revised manuscript, marked in yellow)

Comment 7. The circularly symmetric design allows potential modular stacking, which is interesting. However, additional stacking would increase load. What is the load capacity or payload limit of the robot?

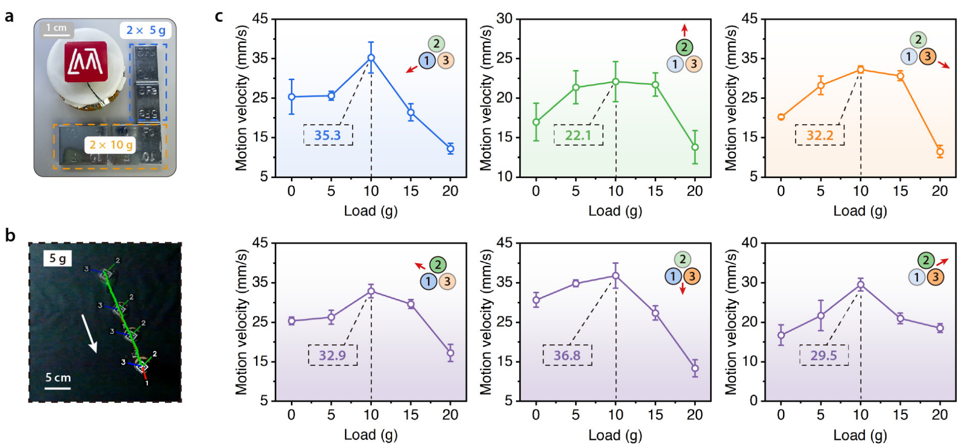
Response: Thank you for this valuable comment.

We have supplemented the manuscript with a series of payload characteristic tests to further evaluate the robot's performance during practical tasks. The experimental results demonstrate that the robot possesses a considerable load-carrying capacity, maintaining stable omnidirectional planar locomotion while supporting a payload of approximately 13.63 times its body weight (robot: 2.2 g; see Part II in **Supplementary Movie S4**). As shown in **Supplementary Fig. 12**, the locomotion speed across all six primary directions follows a non-monotonic trend, characterized by an initial increase followed by a subsequent decline as the payload rises. The peak velocity of 36.8 mm/s (1.19 BL/s) is achieved at an additional load of 10 g under the dual-actuation mode (actuation: #1 and #3; LBS: #2).

This velocity characteristic is primarily attributed to the frictional coupling mechanism between the driving feet and the substrate. At low payloads, the robot's self-weight results in insufficient normal force, leading to partial slippage during the actuation cycle and limiting the efficient transmission of driving energy. As the payload increases to a certain extent, the enhanced friction reduces the slip duration, allowing the driving force to be converted into forward displacement more effectively. However, once the payload exceeds a certain threshold (10 g), the excessive frictional resistance begins to constrain the vibration amplitude of the foot tips, effectively acting as a damping factor that ultimately leads to a decline in locomotion speed. This supplementary evaluation confirms the robot's capability for carrying necessary functional components and highlights the potential for future integration with functional manipulation units.

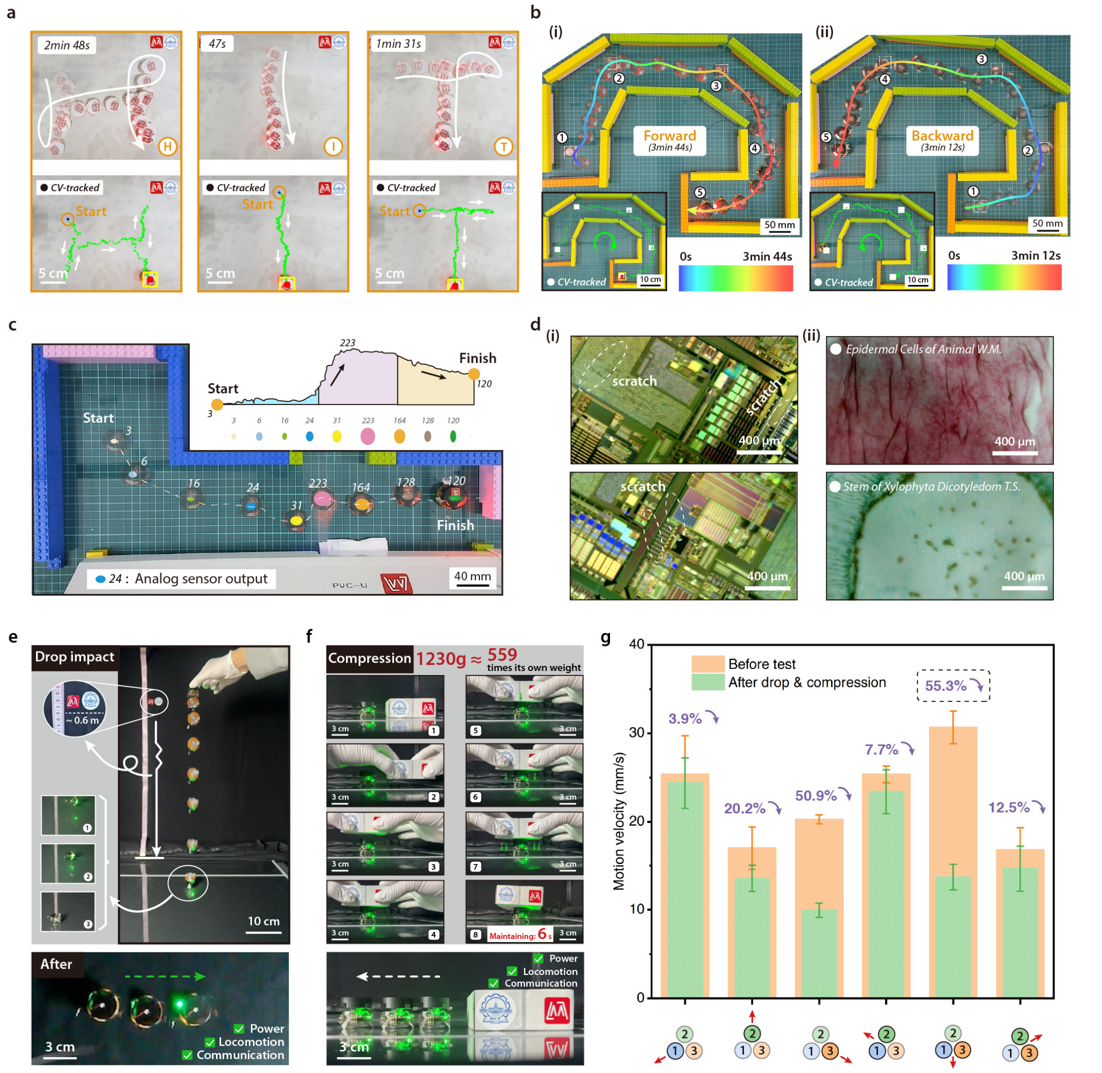
Additionally, we conducted ultimate load-bearing tests with a certain degree of destructiveness to evaluate the structural limits of the robot (**Fig. 4f**). The results demonstrate that the robot can withstand a massive iron block weighing approximately 559 times its own body mass (The original robot: 2.2 g, iron block: 1230 g). Although significant structural deformation occurred under the load for 6 seconds, the robot partially recovered its original shape upon unloading and successfully retained its untethered locomotion capability.

Modification: We have incorporated the experimental results into the revised manuscript.



*"**Supplementary Fig. 12.***

Payload capacity and locomotion characterization. (a)** Photographs of the robot and the standard weights used for testing (5 g and 10 g). **(b)** Time-lapse stacked images showing the robot performing* *locomotion with a 5 g payload.* ***(c) *Locomotion velocity as a function of the applied payload. The robot sustains a maximum additional payload of 20 g. Considering the onboard electronics and battery (10 g), the total carried load reaches 30 g, which is approximately 13.63 times the robot's body weight (the mass of the tripodal robot is 2.2 g). Note: Data obtained from a representative prototype (Prototype #2, 60 V_{p-p}).”* (please see page 29 of the Supplementary Materials, Supplementary Fig. 12, marked in yellow)



*“****Fig 4. Performance and applications for the robot. (a)*** *The untethered robot draws the trajectory of “H”, “I”, and “T” letters on the iron substrate.* ***(b)*** *The robot demonstrates bidirectional maze navigation and reports five magnetic checkpoints it passed using the onboard hall sensor. (i) Forward. (ii) Backward.* ***(c)*** *The gas leak inspection application. (Note: the values only represent the sensor analog values and are not actual gas concentrations)* ***(d)*** *The robot is capable of performing tasks in microscopic observation. (i) Wafer detection. (ii) Cell specimen observation (Epidermal Cells of Animal W.M. and Stem of Xylophyta Dicotyledon T.S.).* ***(e)*** *The robot (Prototype #2) remains functional after a free-fall impact from a height of 0.6 m onto a desktop.* ***(f)*** *The robot (Prototype #2) maintains normal locomotion after sustaining a compressive load of 1230 g (approximately 559 times its body weight).* ***(g)*** *Evaluation of velocity reduction after a drop impact and compressive loading (Prototype #2, 60 V_{p-p}; maximum reduction: 55.3% reduction; Actuation: #1 and #3, LBS: #2).”*

(please see page 37, line 777 of the revised manuscript, Fig 4, marked in yellow)

"Furthermore, the robot exhibits a load-carrying capacity, achieving omnidirectional planar motion (six directions) while carrying a payload of approximately 13.63 times its body weight (robot: 2.2 g; total load: 30 g, including a 10 g power unit and a 20 g additional mass; see Part II in ***Supplementary Movie S4****). As shown in* ***Supplementary Fig. 12****, the locomotion speed in all six directions follows a non-monotonic trend characterized by an initial increase followed by a subsequent decrease as the payload increases. The maximum velocity across six directions is achieved at an additional load of 10 g (peak value: 36.8 mm/s, 1.19 BL/s; actuation: #1 and #3, LBS: #2). This phenomenon is likely attributed to the frictional coupling between the driving feet and the substrate. At lower weights, the robot's self-weight generates insufficient friction to match the driving force, leading to partial slippage during the actuation cycle. An appropriate increase in payload enhances the frictional force, thereby reducing the slip duration and increasing the velocity. However, excessive friction eventually restricts the displacement of the foot tips, resulting in a subsequent decline in locomotion speed. The payload characteristic tests demonstrate the robot's considerable load-carrying capacity, highlighting the potential for future integration with functional manipulation units."*

(please see page 12, line 283 of the revised manuscript, marked in yellow)

"The robot can carry a maximum payload of approximately 13.63 times its body weight and achieves effective omnidirectional locomotion on glass, marble, and stainless-steel surfaces. Furthermore, the robot retains its functional locomotion capabilities after sustaining a 0.6 m drop impact and a continuous compressive load of approximately 559 times its body weight, demonstrating exceptional structural robustness." (please see page 20, line 493 of the revised manuscript, marked in yellow)

Comment 8. Since miniaturization and a small turning radius are critical for operations in confined spaces, it would be very informative to add a figure or table comparing this robot with others reported in the literature in terms of size, turning radius, degrees of freedom, locomotion performance, and power consumption.

Response: We sincerely thank you for this valuable suggestion.

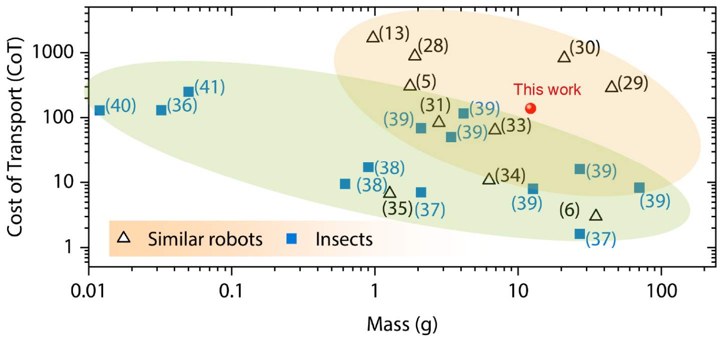
A multi-dimensional comparative table can more intuitively highlight the advantages of the proposed robot. However, as the turning radii vary across different scales, the resulting reorientation costs are difficult to compare directly. To address this, we define a dimensionless metric known as the minimum reorientation cost (RC_{min}) (see **Supplementary Note 1** for a definition). Our findings indicate that the turning radius inherently increases the reorientation cost during a robot's planar locomotion. Thus, we compare the key performance metrics (actuation mode, size, weight, maximum motion speed, minimum motion step distance, RC_{min}, multiple degrees of freedom (DOFs) and endurance) of miniature robots with different actuation types, corresponding to References [5 ~ 17, 20, 21, 28 ~ 35, 42 ~ 44] in the main text, as shown in **Table S1**. The robot combines a miniaturized form factor with a low-cost in-plane reorientation capability and the ability to perform high-resolution movements in an untethered configuration.

In addition, we evaluate the power consumption using the cost of transport (CoT), with a calculated value of 139 (see **Supplementary Note 6**). The CoT is compared with other similar robots (References: 5,6,13,28 ~ 35) and insects (References: 36 ~ 41), as shown in **Supplementary Fig. 18**. The results show that the robot achieves a good balance between high integration and power consumption. The main sources of power consumption are the boost conversion on the circuit board and the power supply requirements of the multiple onboard sensors.

Modification: The comparative table and the power consumption comparison figure have been added to the corresponding locations in the manuscript, with further discussion and analysis incorporated into the Discussion section.

"Furthermore, we evaluate the power consumption using the cost of transport (CoT), with a calculated value of 139 (see ***Supplementary Note 7****). The CoT is compared with other similar robots[5,6,13,28–35] and insects[36–41], as shown in* ***Supplementary Fig. 18****. The results show that the robot achieves a good balance between high integration and power consumption. The main sources of power consumption are the boost conversion on the circuit board and the power supply requirements of the multiple onboard sensors. In addition, we compare the key performance metrics (actuation mode, size, weight, maximum motion speed, minimum motion step distance,* RC_{min}*, multiple degrees of freedom (DOFs) and endurance) of miniature robots with different actuation types[5–17,20,21,28–35,42–44], as shown in* ***Table S1****. The robot combines a miniaturized form factor with a low-cost in-plane reorientation capability and the ability to perform high-resolution movements in an untethered configuration."*

(please see page 21, line 505 of the revised manuscript, marked in yellow)



*"****Supplementary Fig. 18.***

Comparison of the CoT with insects and similar robots. *(Note: Some data are estimated from existing data in their papers.)"*

(please see page 35 of the Supplementary Materials, Supplementary Fig. 18, marked in yellow)

"Table S1. Comparison of key metrics with similar robots.

Actuation mode	Size (cm^3)	Weight (g)	Max. speed (mm/s)/(BL/s)	Min. step (μm)	RC_{min}	DOFs	Endurance
Piezoelectric	Φ3.1×2.75	12.2	30.08 / 0.97	0.56 untethered	1	2	~ 49 min 62.1 m
Piezoelectric[30]	6.3×7.2×1.3	21	0.24 / -	4.06 tethered	1	3	/
Piezoelectric[42]	18.6×18.6×8.7	1282	712 / 3.83	0.49 tethered	1	3	/
Piezoelectric[31,32]	4.5 (length)	2.8	172 / 3.8	~mm scale tethered	>1	2	/
Piezoelectric[28]	2.4 (length)	1.9	28.8 / 1.2	~mm scale untethered	>1	2	~ 19 min
Piezoelectric[43]	4.6×3.6×2.8	34.8	223 / 4.8	0.68 tethered	>1	2	~ 14.5 min
Piezoelectric[29]	5.5×5.5×1.4	45	31.5 / 0.57	~mm scale tethered	1	3	/
Piezoelectric[33]	3×3×1.43	6.9	255 / 8.5	0.25 tethered	>1	2	/
Piezoelectric[44]	2.5×2.5×2.6	1.74	120 / 4.8	~mm scale tethered	>1	2	/
Piezoelectric[34]	5×1×0.15	6.27	302 / 6.04	~mm scale tethered	-	1	/
Piezoelectric[35]	4.35 (length)	1.27	440 / 10.1	~mm scale tethered	>1	2	/
DC motor[5]	2 (length)	1.76	350 / 17.5	~mm scale untethered	>1	2	/
DC motor[6]	7.5×9.5×2.1	35	815 / 10.9	~mm scale untethered	>1	2	~ 25 min
DC motor[7]	10 (length)	16	1500 / 15	~mm scale untethered	>1	2	/
DC motor[8]	6.4 (length)	/	320 / 5	~mm scale tethered	>1	2	/
DC motor[9]	2.5 (length)	1.2	28.3 / 1.1	~mm scale untethered	>1	2	/
Dielectric elastomer[10]	9.1 (length)	/	124 / 1.36 (kV level)	~mm scale tethered	>1	2	/
Dielectric elastomer[11]	10 (length)	9.6	78 / 0.78 (kV level)	~mm scale tethered	>1	2	/
Dielectric elastomer[12]	/	~ 0.62	- / 20.6 (kV level)	~mm scale tethered	-	1	/
Dielectric elastomer[13]	4 (length)	0.97	12 / 0.3 (~ 500V level)	~mm scale untethered	>1	2	~ 15 min

Magnetic[14]	0.9 (length)	0.71	18.9 / 2.1	~mm scale untethered	>1	2	~ 4.5 min
Magnetic[15]	1.5 (length)	0.0455	8 / 0.53	~mm scale limited, magnetic field	-	2	/
Magnetic[16]	0.3 (length)	0.005	213 / -	~mm scale limited, magnetic field	-	2	/
Magnetic[17]	0.8 (length)	/	97.5 / 12.2	~mm scale limited, magnetic field	-	2	/
Shape memory alloy[20]	4.2×0.6×0.4	0.61	0.15 / -	- limited, heating	-	1	/
Shape memory alloy[21]	3 (length)	2.4	30 / 1	~mm scale untethered	>1	2	/

(Note: Some data are estimated from existing data in their paper; priority is given to comparisons involving untethered locomotion data. The degrees of freedom (DOFs) metric primarily focuses on the capability for planar translation along the X and Y axes and rotation about the Z-axis. Notably, single-DOF robots and robots dependent on external magnetic fields are excluded from the RC_{min} evaluation.)"
(please see page 37 of the Supplementary Materials, Table S1, marked in yellow)

Comment 9. The demonstration of multi-robot position-data sharing is impressive. Showing collaboration among three or more robots would make the scenario even more compelling.

Response: We sincerely thank you for this constructive suggestion.

We agree that demonstrating collaboration among a larger group of robots can strengthen the evidence for the scalability and practical utility of our proposed navigation system. In the revised manuscript, we have added new experimental results involving a three-robot system operating on a robot vision feedback platform (see Part II in **Supplementary Movie S12**). This demonstration highlights the PTP navigation system's ability to handle multiple agents and its adaptability to different architectural frameworks.

Multi-robot motion demonstration is achieved by integrating the vision feedback platform with the PTP navigation algorithm. Initially, we consider a single-robot scenario where the user can select the target robot within the video stream display area and designate a target position (**Supplementary Fig. 17a**). The host computer then transmits the target coordinates to the robot. Leveraging the onboard PTP navigation system, the robot autonomously calculates the required actuation ratios for its driving legs, enabling point-to-point motion without the need for users to manage orientation deflection or complex control commands. In a multi-robot scenario, three robots (ID: 0, 3, and 4) are selected as the controlled agents, with robot ID: 0 designated as the "Leader." Upon the user specifying a target point, the leader (ID: 0) autonomously navigates toward the goal, while the remaining robots (Followers) autonomously follow the leader (**Supplementary Fig. 17b**).

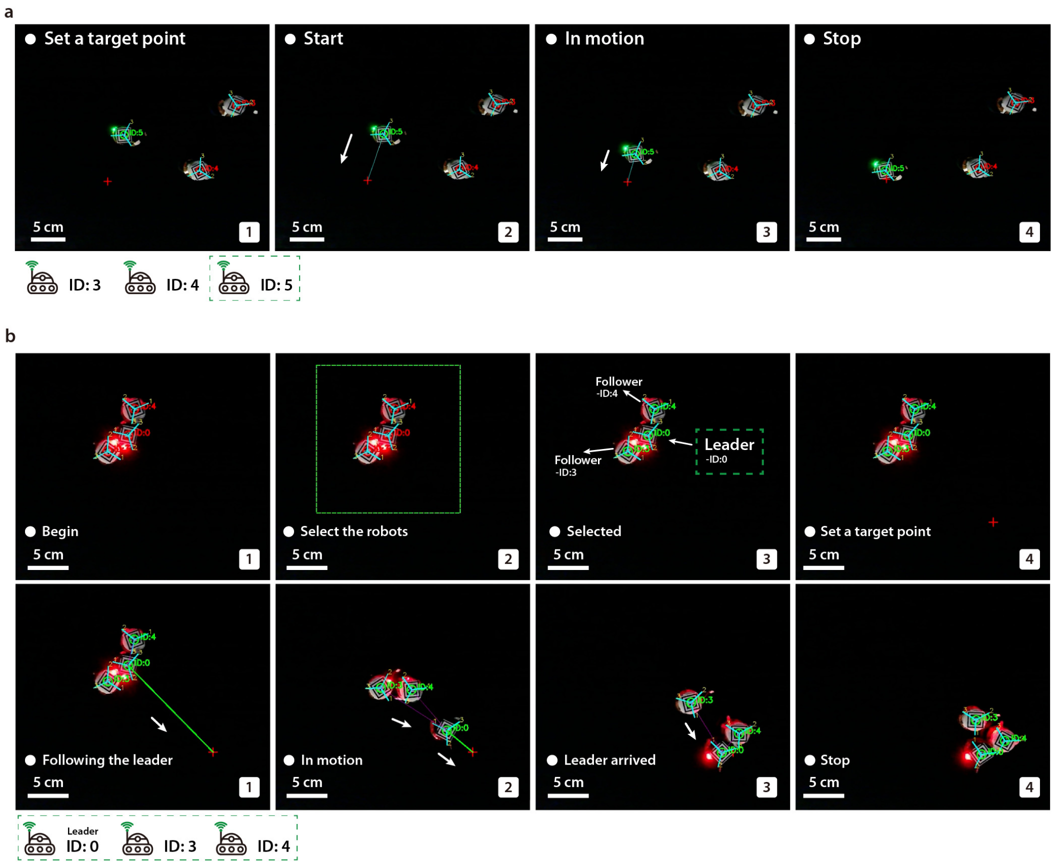
Modification: We have incorporated the experimental results and descriptions into the revised manuscript.

"Furthermore, we increase the number of robots to three and transition to a robot vision feedback platform, demonstrating that the PTP navigation system exhibits good scalability in terms of both the number of controlled agents and the underlying platform framework (see Part II in ***Supplementary Movie S12****). Multi-robot motion demonstration is achieved by integrating the vision feedback platform with the PTP navigation algorithm. Initially, we consider a single-robot scenario where the user can select the target robot within the video stream display area and designate a target position (****Supplementary Fig. 17a****). The host computer then transmits the target coordinates to the robot. Leveraging the onboard PTP navigation system, the robot autonomously calculates the required actuation ratios for its driving legs, enabling point-to-point motion without the need for users to manage orientation deflection or complex control commands. In a multi-robot scenario, three robots (ID: 0, 3, and 4) are selected as the controlled agents, with robot ID: 0 designated as the "Leader." Upon the user specifying a target point, the leader (ID: 0) autonomously navigates toward the goal, while the remaining robots (Followers) autonomously follow the leader (****Supplementary Fig. 17b****). The entire group ceases motion once all following robots have reached the target destination."*

(please see page 19, line 462 of the revised manuscript, marked in yellow)

"We implement three multi-robot auto-following scenarios, demonstrating the scalability and practical application potential of the PTP navigation system."

(please see page 21, line 503 of the revised manuscript, marked in yellow)



"Supplementary Fig. 17.

The PTP navigation system for multi-robot following applications. (a) *Users can control the motion of* *a single robot through the interactive interface of the video stream based on the robot vision feedback platform.* ***(b)*** *Two follower robots track the leader robot (ID: 0) to achieve coordinated motion using the PTP navigation algorithm."*

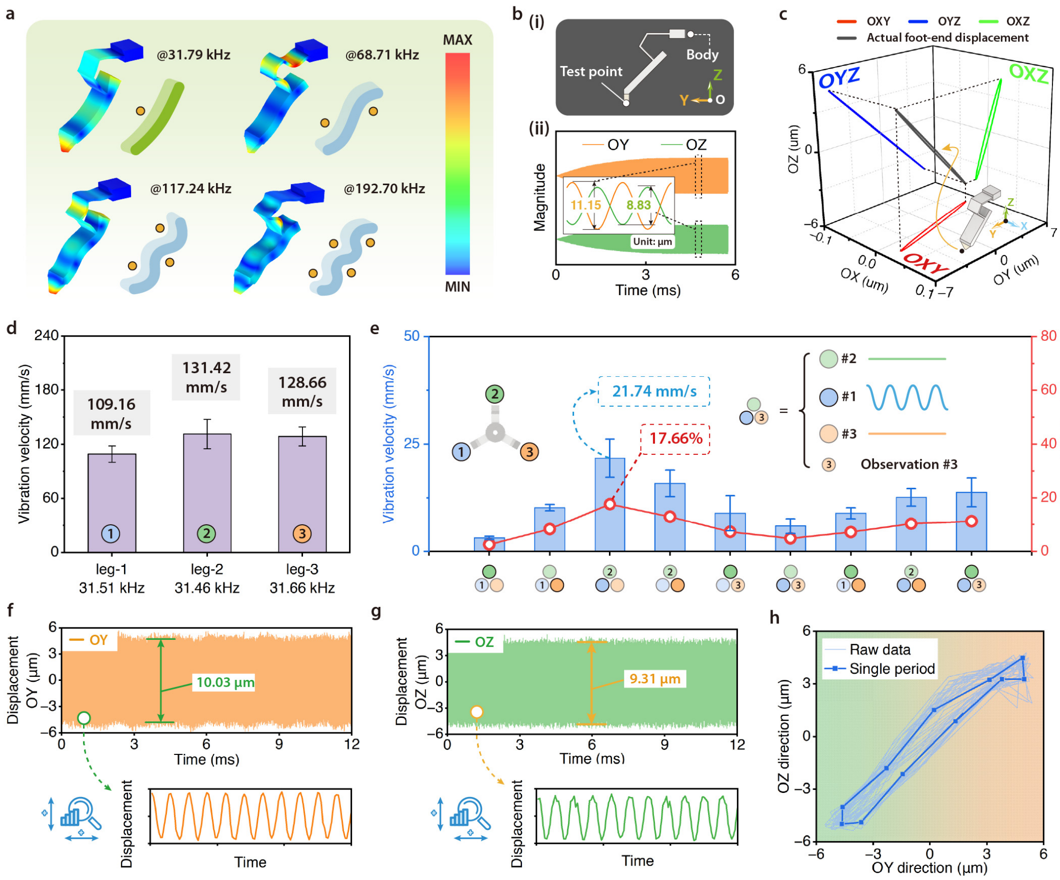
(please see page 34 of the Supplementary Materials, Supplementary Fig. 17, marked in yellow)

Comment 10. Some figures miss data labels, such as Fig. 2c and Fig. 2h; these should be clearly marked.

Response: We thank you for pointing out this oversight.

We have added the missing labels to **Fig. 2c** and **Fig. 2h**. Furthermore, we conducted a thorough review of all figures for similar issues and rectified them in **Fig. 2f** and **Fig. 2g**.

Modification: **Figure 2** in the main text has been revised accordingly.



*"**Fig 2. Robot simulation analysis and vibration characteristics. (a)** Bending modes (1 ~ 4) and resonant frequencies of the actuation unit. **(b)** Transient dynamics simulation results of the robot's foot end. (i) The measurement point is at the foot end. Note: the other two driving legs are hidden for clarity but are included in the simulation. (ii) Simulation results: OY: 11.15 μm, OZ: 8.83 μm, 80 $V_{p\text{-}p}$. **(c)** Spatial trajectory of the robot's foot end. **(d)** Vibration velocity characteristics of driving legs (10 $V_{p\text{-}p}$, leg-1: 109.16 mm/s, leg-2: 131.42 mm/s, leg-3: 128.66 mm/s, fixed frequency mode). Error bars: three experimental trials. **(e)** The vibration isolation characteristic test (10 $V_{p\text{-}p}$, maximum vibration velocity: 21.74 mm/s, corresponding transmissibility: 17.66%, actuation: #1, observation: #2, fixed frequency mode). Colored circles with 100% and 50% transparency represent the actuation unit under OVS and the non-actuated units, respectively. The numbers on the colored circles indicate the observed actuation unit. Note: the excitation frequency is taken as the mean of*

their respective resonant frequencies when the two actuation units are required to operate simultaneously. ***(f)*** *and* ***(g)*** *show the vibration displacements of the prototype foot end in the OY and OZ directions, respectively (OY: 10.03 μm, OZ: 9.31 μm, 80* $V_{p\text{-}p}$*).* ***(h)*** *The diagonal vibration trajectory of the foot end. The light blue and dark blue curves represent the raw data and the complete trajectory over a single period, respectively."*

(please see page 33, line 744 of the revised manuscript, marked in yellow)

Responses to reviewers' comments (Reviewer: #2)

Comments to the Author:

This paper presents the design and experimental validation of a miniature legged robot driven by piezoelectric actuators. A key feature of the robot is its ability to move omnidirectionally without turning, achieved by selectively actuating one or two of its legs. The system is interesting and demonstrates a high level of integration among onboard sensors, actuators, and power components. The authors also provide several compelling demonstrations. Overall, the paper is well written and easy to follow. However, the reviewer has several comments that may help strengthen the paper:

Response: We sincerely thank you for your efforts in reviewing this article, your approval and valuable suggestions. Your professional comments are very helpful for the improvement of this article and the direction of our follow-up work. The manuscript has been carefully revised according to your suggestions. A point-by-point response is provided below, and all corresponding modifications have been highlighted in yellow in the revised manuscript and Supplementary Information. We hope these revisions satisfactorily address your concerns.

Comment 1. The authors emphasize the robot's ability to move without turning. However, it is unclear why minimizing the turning radius is critical for such a small robot. If the robot is already miniature in size, having a small turning radius might not pose a significant limitation. Please elaborate on the practical importance of this design goal.

Response: We thank you for your professional comment.

1. We agree with your perspective that a robot with a miniaturized footprint holds great promise for navigating restricted spaces, even with a finite turning radius. However, it should be noted that comparisons are only meaningful when conducted across the same dimensions. At a comparable scale, minimizing the turning radius can further enhance a robot's mobility and practical utility.

2. In the introduction, we pointed out that the design of centimeter-scale robots increasingly emphasizes miniaturization and high system integration. These two conflicting objectives force a fundamental trade-off: maximizing onboard functionality while minimizing physical footprint. By analyzing the strengths and limitations of typical microrobots based on various actuation principles, we found that robots based on piezoelectric actuators offer unique advantages for miniaturization and multi-functional integration. However, in the process of achieving this, we discovered that a robot's locomotion performance depends not only on its size but also on its turning radius. We wish to emphasize that while utilizing the novel smart actuation principle of piezoelectric actuation represents an attempt to tackle the challenges of structural miniaturization, it does not provide a way to eliminate the locomotion costs (space and time) incurred by the turning radius.

A robot's maneuverability is not only defined by its size but also by its turning radius. A robot might not be able to make turns in a space where it can easily pass. To address this, we define a dimensionless metric known as the minimum reorientation cost (RC_{min}) (see **Supplementary Note 1** for a definition). Our findings indicate that the turning radius inherently increases the reorientation cost during a robot's planar locomotion. In this context, the inevitable turning radius of microrobots driven by piezoelectric actuators imposes comparable constraints on spatial reorientation costs, thereby limiting their accessibility to confined environments. It is because most current reports on miniature piezoelectric robots use linear decoupling drive methods. Robots using this method typically possess a single linear driving force (forward and backward) and rely on reorienting their own attitude to change the direction of actuation covering the full range (0 ~ 360°). Frequent attitude adjustments and re-positioning within confined spaces further complicate the task of motion control. Previous works have partially addressed this challenge through two main strategies. The first aims to reduce the turning radius mechanically. For example, some miniature robots utilize electrostatic footpads to create a pivot foot, effectively reducing the turning radius to the robot's width. The second strategy relies on tangentially distributed force synthesis for omnidirectional motion. Although this configuration eliminates the turning radius, it is inherently optimized for generating rotational torque, which makes achieving stable linear motion difficult due to its susceptibility to undesirable yaw and path deviation.

Therefore, it remains a significant challenge to leverage piezoelectric actuation for robot miniaturization to facilitate highly integrated multi-sensor packaging, while simultaneously enabling planar locomotion without attitude adjustments to eliminate the reorientation costs associated with a turning radius. Addressing this challenge is precisely where the significance of our research lies.

3. Our robot achieves a high level of integration—incorporating actuation, control, communication, power supply, and five types of onboard sensors—within a centimeter-scale body (3 cm in diameter). With further integration of sensors and essential driving circuits, the size could be even further reduced. Building on this compact design, we employ a tripodal radial actuation strategy and a trajectory interpolation algorithm to achieve planar locomotion without attitude turning. We compare the key performance metrics (actuation mode, size, weight, maximum motion speed, minimum motion step distance, $RC_{\min}$, multiple degrees of freedom (DOFs) and endurance) of miniature robots with different actuation types (5 ~ 17, 20, 21, 28 ~ 35, 42 ~ 44), as shown in **Table S1**. The robot combines a miniaturized form factor with a low-cost in-plane reorientation capability and the ability to perform high-resolution movements in an untethered configuration.

Modification: We have incorporated the relevant descriptions into the Introduction section to enhance its logical coherence, and provided a detailed analysis for the minimum reorientation cost ($RC_{\min}$) in the supplementary materials.

*"A robot's maneuverability is not only defined by its size but also by its turning radius. A robot might not be able to make turns in a space where it can easily pass. To address this, we define a dimensionless metric known as the minimum reorientation cost ($RC_{\min}$) (see **Supplementary Note 1** for a definition). Our findings indicate that the turning radius inherently increases the reorientation cost during a robot's planar locomotion."* (please see page 3, line 53 of the revised manuscript, marked in yellow)

"Supplementary Note 1

Minimum reorientation cost

Due to operational requirements such as obstacle avoidance or task execution, the planar locomotion of robots inevitably involves a sequence of reorientation processes. The 90° reorientation processes for both an omnidirectional robot and a robot with a finite turning radius are compared in Supplementary Fig. 1. Robots with a finite turning radius must adjust their orientation before moving toward the target, whereas omnidirectional robots can proceed directly to the target point. To quantitatively evaluate reorientation costs, we define a dimensionless metric termed the minimum reorientation cost ($RC_{\min}$), representing the cost incurred by the robot during a single reorientation of $\pi/2$. We define A_{body} as the area of the maximum *circumscribed circle of the robot's vertical projection onto the plane. Consequently, the $RC_{\min}$ is defined as the ratio of the minimum area swept by A_{body} (denoted as A_{swept}) during a single reorientation of π/2 to A_{body} itself. That is:*

$$RC_{\min} = \frac{A_{\text{swept}}}{A_{\text{body}}} = \frac{A_{\text{swept}}}{\pi {R_{\max}}^2} \qquad (15)$$

Where $R_{\max}$ is the radius of the maximum circumscribed circle of the robot's planar projection (Unit: mm).

*For omnidirectional robots (OM robots), the transition from a forward heading to a lateral leftward trajectory involves no attitude adjustment and is achieved solely by altering the drive signals (see **Supplementary Fig. 1a**). Consequently, the area swept during reorientation remains A_{body}, yielding an $RC_{\min}$ of 1.*

$$RC_{\min}^{\text{OM}} = 1 \qquad (16)$$

*For a robot with a finite turning radius (TR robots), motion is constrained by a turning radius r_{turn} (assuming a constant radius during the steering maneuver). The reorientation process is illustrated in **Supplementary Fig. 1b**. Consequently, the $RC_{\min}$ is given by:*

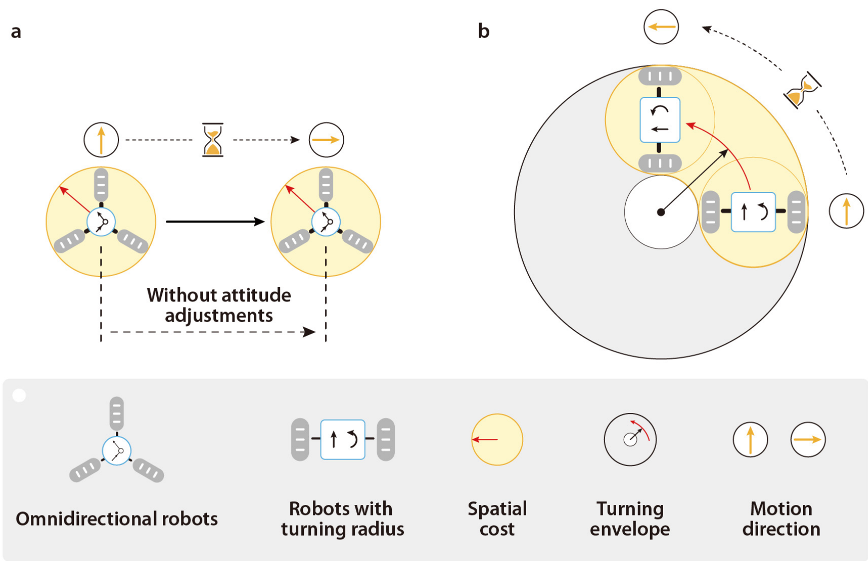
$$RC_{\min}^{\text{TR}} = \frac{\frac{\pi/2}{2}[(r_{\text{turn}} + R_{\max})^2 - (r_{\text{turn}} - R_{\max})^2] + \pi {R_{\max}}^2}{\pi {R_{\max}}^2} = 1 + \frac{r_{\text{turn}}}{R_{\max}} \qquad (17)$$

In conclusion, the ranking of $RC_{\min}$ for the two robot categories is:

$$RC_{\min}^{\text{OM}} < RC_{\min}^{\text{TR}} \qquad (18)$$

Thus, the omnidirectional motion robot exhibits the lowest reorientation cost, while the TR robot incurs the higher cost. This advantage affords the OM robot better traversability in narrow spaces and reduced control overhead. The simplification of control logic is particularly critical in scenarios involving high-frequency, computationally intensive operations, such as closed-loop navigation and path planning. Admittedly, this cost estimation method implies certain limitations. For instance, it primarily focuses on the angular cost associated with the robot's self-reorientation, while the operational energy expenditure throughout the entire trajectory and the computational complexity of the actual motion algorithms are not fully accounted for."

(please see page 3 of the Supplementary Materials, Supplementary Note 1, marked in yellow)



"Supplementary Fig. 1.
The calculation of the minimum reorientation cost. (a) *RC_{min} of the omnidirectional robots.* ***(b)*** *RC_{min} of the robots with turning radius."*

(please see page 18 of the Supplementary Materials, Supplementary Fig. 1, marked in yellow)

"Table S1. Comparison of key metrics with similar robots.

Actuation mode	Size (cm^3)	Weight (g)	Max. speed (mm/s)/(BL/s)	Min. step (μm)	RC_{min}	DOFs	Endurance
Piezoelectric	Φ3.1×2.75	12.2	30.08 / 0.97	0.56 untethered	1	2	~ 49 min 62.1 m
Piezoelectric[30]	6.3×7.2×1.3	21	0.24 / -	4.06 tethered	1	3	/
Piezoelectric[42]	18.6×18.6×8.7	1282	712 / 3.83	0.49 tethered	1	3	/
Piezoelectric[31,32]	4.5 (length)	2.8	172 / 3.8	~mm scale tethered	>1	2	/
Piezoelectric[28]	2.4 (length)	1.9	28.8 / 1.2	~mm scale untethered	>1	2	~ 19 min
Piezoelectric[43]	4.6×3.6×2.8	34.8	223 / 4.8	0.68 tethered	>1	2	~ 14.5 min
Piezoelectric[29]	5.5×5.5×1.4	45	31.5 / 0.57	~mm scale tethered	1	3	/
Piezoelectric[33]	3×3×1.43	6.9	255 / 8.5	0.25 tethered	>1	2	/
Piezoelectric[44]	2.5×2.5×2.6	1.74	120 / 4.8	~mm scale tethered	>1	2	/
Piezoelectric[34]	5×1×0.15	6.27	302 / 6.04	~mm scale tethered	-	1	/
Piezoelectric[35]	4.35 (length)	1.27	440 / 10.1	~mm scale tethered	>1	2	/

DC motor[5]	2 (length)	1.76	350 / 17.5	~mm scale untethered	>1	2	/
DC motor[6]	7.5×9.5×2.1	35	815 / 10.9	~mm scale untethered	>1	2	~ 25 min
DC motor[7]	10 (length)	16	1500 / 15	~mm scale untethered	>1	2	/
DC motor[8]	6.4 (length)	/	320 / 5	~mm scale tethered	>1	2	/
DC motor[9]	2.5 (length)	1.2	28.3 / 1.1	~mm scale untethered	>1	2	/
Dielectric elastomer[10]	9.1 (length)	/	124 / 1.36 (kV level)	~mm scale tethered	>1	2	/
Dielectric elastomer[11]	10 (length)	9.6	78 / 0.78 (kV level)	~mm scale tethered	>1	2	/
Dielectric elastomer[12]	/	~ 0.62	- / 20.6 (kV level)	~mm scale tethered	-	1	/
Dielectric elastomer[13]	4 (length)	0.97	12 / 0.3 (~ 500V level)	~mm scale untethered	>1	2	~ 15 min
Magnetic[14]	0.9 (length)	0.71	18.9 / 2.1	~mm scale untethered	>1	2	~ 4.5 min
Magnetic[15]	1.5 (length)	0.0455	8 / 0.53	~mm scale limited, magnetic field	-	2	/
Magnetic[16]	0.3 (length)	0.005	213 / -	~mm scale limited, magnetic field	-	2	/
Magnetic[17]	0.8 (length)	/	97.5 / 12.2	~mm scale limited, magnetic field	-	2	/
Shape memory alloy[20]	4.2×0.6×0.4	0.61	0.15 / -	- limited, heating	-	1	/
Shape memory alloy[21]	3 (length)	2.4	30 / 1	~mm scale untethered	>1	2	/

(Note: Some data are estimated from existing data in their paper; priority is given to comparisons involving untethered locomotion data. The degrees of freedom (DOFs) metric primarily focuses on the capability for planar translation along the X and Y axes and rotation about the Z-axis. Notably, single-DOF robots and robots dependent on external magnetic fields are excluded from the RC_{min} evaluation.)"

(please see page 37 of the Supplementary Materials, Table S1, marked in yellow)

Comment 2. While the robot is claimed to achieve both micro- and macro-scale motion, the experimental videos show that its motion is not very straight. The trajectory appears to deviate, suggesting potential limitations in movement accuracy. Can the robot reliably follow a straight line? Please comment on the control or mechanical factors that contribute to this behavior.

Response: We appreciate your professional comment.

It is true that because the vibration amplitude of the driving feet is at the micrometer scale, the robot is inherently sensitive to external disturbances (such as microscopic obstacles or surface irregularities), which can lead to unpredictable trajectory deviations or attitude yaw. Beyond these external factors, several internal factors contribute to the observed inconsistency in locomotion characteristics:

1. Mechanical manufacturing and assembly errors

Manufacturing and installation errors between the legs and the center block can induce parasitic torques, causing unintentional rotation during linear motion. In addition, microscopic differences in foot-ground contact and uneven wear over time contribute to frictional imbalances, affecting overall path stability.

2. Accumulative errors in open-loop control and motion synthesis

As the robot currently operates under open-loop control with a discrete motion resolution (12 directions), deviations primarily arise from the cumulative errors during trajectory interpolation and force synthesis.

Thus, we developed a point-to-point (PTP) navigation system (**Fig. 5a-b**) that allows the robot to independently handle accidental deflections. Leveraging the onboard PTP navigation system, the robot autonomously calculates the required actuation ratios for its driving legs, enabling seamless point-to-point motion without the need for users to manage orientation deflection or complex control commands.

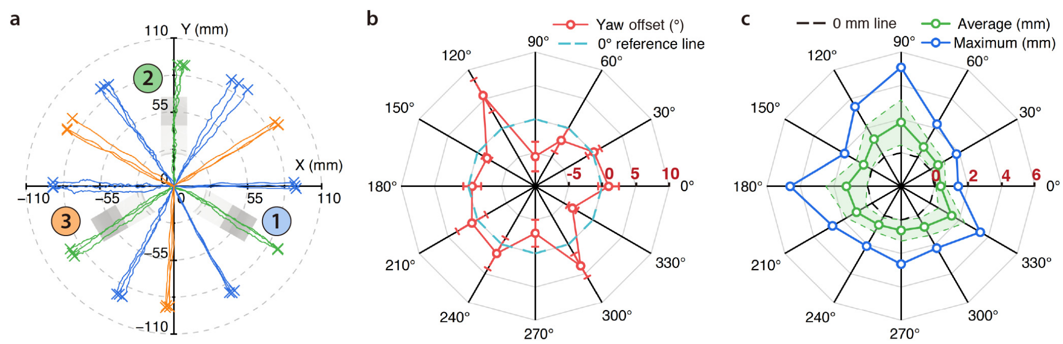
In addition, to further evaluate the reliability in maintaining a straight trajectory, we conducted additional tests to evaluate the robot's omnidirectional synthetic motion capabilities across 12 directions with 30° increments (see Part III in **Supplementary Movie S3**). For each direction, a 90-mm trajectory segment is extracted for analysis (**Supplementary Fig. 10a**). Orientation stability is evaluated via yaw offset, defined as the angular difference between the theoretical target and the actual trajectory angle (ranging from 0° to 360° relative to the positive x-axis). The orientation stability is evaluated via yaw offset; the results show a minimum yaw offset of 0.2° during motion toward the 30° azimuth, while a maximum offset of 5.6° occurred at the 120° azimuth (**Supplementary Fig. 10b**). Furthermore, the trajectory deviation, including the average and maximum trajectory deviations, are employed to quantify trajectory-keeping capabilities (definitions provided in **Supplementary Fig. 10c** caption). The average deviation ranged from a minimum of 0.4 mm (0.4% relative to travel distance) at 0° to a maximum of 1.8 mm (2.0%) at 90°, while the maximum deviation spanned from 1.4 mm (1.6%) at 0° to 5.1 mm (5.7%) at 90°. These results demonstrate that the robot, governed by the tripodal radial actuation strategy, effectively achieves planar motion across 12 discretized directions (30° resolution) with trajectory-keeping capabilities.

In future research, we try to develop a mathematical model coupling the robot's vibration and friction characteristics to more clearly reveal the underlying physics of its dynamic locomotion. Additionally, we will explore more efficient vibration isolation methods and implement more robust closed-loop control algorithms for full-trajectory tracking.

Modification: We have added relevant descriptions and analyses to the revised manuscript.

*"By modulating the driving force ratios of the two active legs in OVS, we find that the tripodal radial actuation strategy can also synthesize six additional asymmetric motion angles. We characterize the robot's omnidirectional motion across 12 directions in 30° increments, including six asymmetric synthetic angles (see Part III in **Supplementary Movie S3**). For each direction, a 90-mm trajectory segment is extracted for analysis (**Supplementary Fig. 10a**). Orientation stability is evaluated via yaw offset, defined as the angular difference between the theoretical target and the actual trajectory angle (ranging from 0° to 360° relative to the positive x-axis). The results show a minimum yaw offset of 0.2° during motion toward the 30° azimuth, while a maximum offset of 5.6° occurred at the 120° azimuth (**Supplementary Fig. 10b**). Furthermore, the trajectory deviation, including the average and maximum trajectory deviations, are employed to quantify trajectory-keeping capabilities (definitions provided in **Supplementary Fig. 10c** caption). The average deviation ranges from a minimum of 0.4 mm (0.4% relative to travel distance) at 0° to a maximum of 1.8 mm (2.0%) at 90°, while the maximum deviation spans from 1.4 mm (1.6%) at 0° to 5.1 mm (5.7%) at 90°. These results demonstrate that the robot, governed by the tripodal radial actuation strategy, effectively achieves planar motion across 12 discretized directions (30° resolution) with trajectory-keeping capabilities."*

(please see page 11, line 260 of the revised manuscript, marked in yellow)



*"**Supplementary Fig. 10.***

***Demonstration and evaluation of planar motion across twelve discrete directions. (a)** Trajectory map of the robot across 12 directions with 30° increments (each movement spans 90 mm; $n = 3$ trials per direction). **(b)** Yaw offset across different motion directions. The yaw angle is defined as the orientation of the motion trajectory relative to the positive x-axis (ranging from 0° to 360°), and the offset represents the angular difference between the actual trajectory and the theoretical target angle. Minimum value: offset 0.2° at 30°, maximum value: offset 5.6° at 120°. **(c)** Evaluation of trajectory deviations across motion directions. For a 90-mm trajectory segment, the average deviation ranges from a minimum of 0.4 mm (0.4% relative to travel distance) at 0° to a maximum of 1.8 mm (2.0%) at 90°. The maximum deviation ranges from a minimum of 1.4 mm (1.6%) at 0° to a maximum of 5.1 mm (5.7%) at 90°. The green dashed area represents the standard deviation (SD). **Note:** Data obtained from a representative prototype (Prototype #2). The green dashed area represents the standard deviation (SD). **Note: Deviation metrics.** Trajectory precision is quantified by the average deviation ($D_{\text{avg}} = \frac{1}{n}\sum |d_i|$) and maximum deviation ($D_{\max} = \max|d_i|$). Here, d_i represents the perpendicular distance of each actual trajectory point from the straight line connecting the segment's start and end points."*

(please see page 27 of the Supplementary Materials, Supplementary Fig. 10, marked in yellow)

"iv) Manufacturing and installation errors induce parasitic torques, which, combined with the constraints in motion synthesis and trajectory interpolation precision, inevitably lead to deviations in the robot's path."

(please see page 22, line 526 of the revised manuscript, marked in yellow)

Comment 3. Since the robot's motion is vibration-based, it appears to require a smooth surface for locomotion. How would it perform on textured or uneven surfaces? Even small obstacles might prevent movement. Please discuss the robot's limitations and potential strategies for improving robustness on non-smooth terrains.

Response: Thank you for your professional advices.

We agree with your analysis regarding the robot's performance on uneven terrains. The locomotion mechanism of our tripodal robot is based on the first-order bending vibration of the piezoelectric actuators, which excites a high-frequency, in-plane oblique trajectory at the foot tips. As you correctly noted, the vibration amplitudes of the robot's foot tips are typically on the micrometer scale (OY: 11.15 μm, OZ: 8.83 μm; see **Fig. 2b** in the main text). This type of movement relies on the high-frequency accumulation of microscopic steps to achieve macroscopic speed. While this strategy enables high-precision motion and rapid response on flat substrates (e.g., glass, marble, or stainless steel), such minute displacements are inherently insufficient to overcome the topographical obstacles of uneven surfaces, such as gravel terrain, where the surface roughness often exceeds the vibration amplitude. This is indeed a common characteristic of vibration-driven piezoelectric robots designed for precision robotics. Our work is primarily situated in the context of applications like micro-manipulation or inspections within controlled environments, where operational surfaces are relatively smooth. Nevertheless, we recognize that enhancing the adaptability to rougher or heterogeneous surfaces is a critical frontier for expanding the application scope of piezoelectric robots.

To enhance the robustness and terrain adaptability of piezoelectric robots in future research, several strategies could be explored. On one hand, a promising approach involves investigating novel vibration modes with larger amplitudes, integrated with optimized driving leg structures. On the other hand, the integration auxiliary wheels could improve frictional coupling and shock absorption on heterogeneous surfaces, ensuring stable motion even under topographical irregularities.

Beyond the glass substrate, the robot also demonstrates effective locomotion on marble and stainless-steel surfaces (see Part I in **Supplementary Movie S4**). The velocity characteristics across the six primary motion directions on these diverse contact interfaces are presented in **Supplementary Fig. 11**. The robot achieves a peak velocity of 30.66 mm/s on the glass substrate (actuation: #1 and #3; low-frequency bounce state (LBS): #2). In comparison, the maximum locomotion speeds on marble and stainless-steel surfaces reach 26.21 mm/s and 26.06 mm/s, respectively (actuation: #1 and #2; LBS: #3 for both cases).

We have performed additional robustness tests, including free-fall impact and compression tests, to validate the robot's survivability and structural integrity under extreme conditions. As shown in **Fig. 4e**, the robot is elevated to a height of 0.6 m and released in its normal working orientation for a free-fall impact onto the working surface. Despite flipping and rolling upon impact, the robot maintained normal power supply, communication, and locomotion functionality (see Part I in **Supplementary Movie S10**). Subsequently, a massive iron block, weighing approximately 559 times (1230 g) the original tripodal robot (2.2 g), is placed on the robot for 6 seconds (**Fig. 4f**). Although structural deformation occurred under the load, the robot partially recovers its original shape after the load is removed and retain its capability for untethered locomotion (see Part II in **Supplementary Movie S10**).

The velocity characteristics across six primary directions are evaluated (**Fig. 4g**). While the robot preserves its locomotion capability in all directions, varying degrees of speed reduction are observed. The maximum reduction occurs in the synthetic direction of driving legs #1 and #3 (dual actuation mode), representing a decrease of approximately 55.3%. This is followed by the single actuation mode (actuation: #3; LBS: #1 and #2), which shows a 50.9% decline. These results suggest that driving leg #3 sustains relatively more severe damage during the two tests, which alters the overall vibration characteristics.

Modification: We have addressed the discussion on locomotion surface limitations and incorporated the supplementary robustness experiments into the corresponding sections of the revised manuscript.

"Beyond the glass substrate, the robot also demonstrates effective locomotion on marble and stainless-steel surfaces (see Part I in ***Supplementary Movie S4****). The velocity characteristics across the six primary motion directions on these diverse contact interfaces are presented in* ***Supplementary Fig. 11****. The robot achieves a peak velocity of 30.66 mm/s on the glass substrate (actuation: #1 and #3; LBS: #2). In comparison, the maximum locomotion speeds on marble and stainless-steel surfaces reach 26.21 mm/s and 26.06 mm/s, respectively (actuation: #1 and #2; LBS: #3 for both cases)."*

(please see page 12, line 277 of the revised manuscript, marked in yellow)

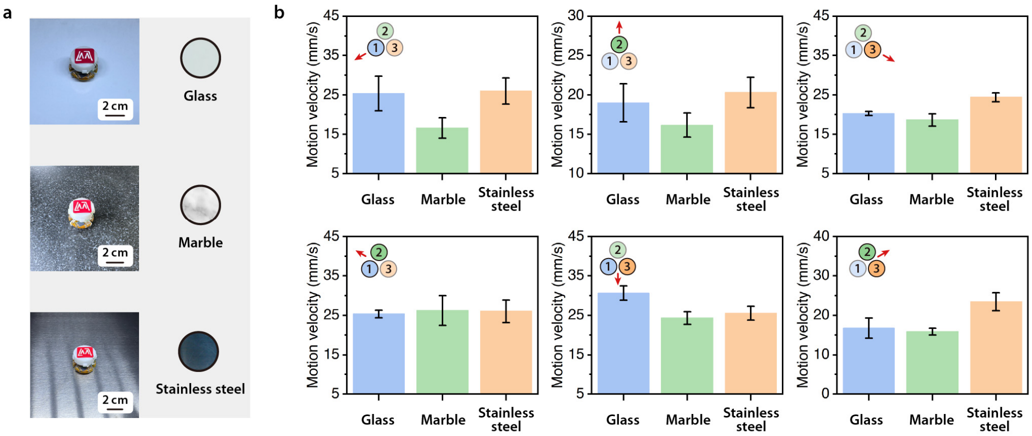
*"To evaluate the robustness required for practical applications, the robot is elevated to a height of 0.6 m and released in its working posture for a free-fall impact onto the working surface (****Fig. 4e****). Despite flipping and rolling upon impact, the robot maintained normal power supply, communication, and locomotion functionality (see Part I in* ***Supplementary Movie S10****). Subsequently, a massive iron block, weighing approximately 559 times (1230 g) the original tripodal robot (2.2 g), is placed on the robot for 6 seconds (****Fig. 4f****). Although structural deformation occurred under the load, the robot partially recovers its original shape after the load is removed and retain its capability for untethered locomotion (see Part II in* ***Supplementary Movie S10****). Following the two sets of experiments, the degradation of velocity characteristics across the six primary directions is evaluated (****Fig. 4g****). While the robot preserves its locomotion capability in all directions, varying degrees of speed reduction are observed. The maximum reduction occurs in the synthetic direction of driving legs #1 and #3 (dual actuation mode), representing a decrease of approximately 55.3%. This is followed by the single actuation mode (actuation: #3; LBS: #1 and #2), which shows a 50.9% decline. These results suggest that driving leg #3 sustains relatively more severe damage during the two tests, which alters the overall vibration characteristics."*

(please see page 16, line 381 of the revised manuscript, marked in yellow)

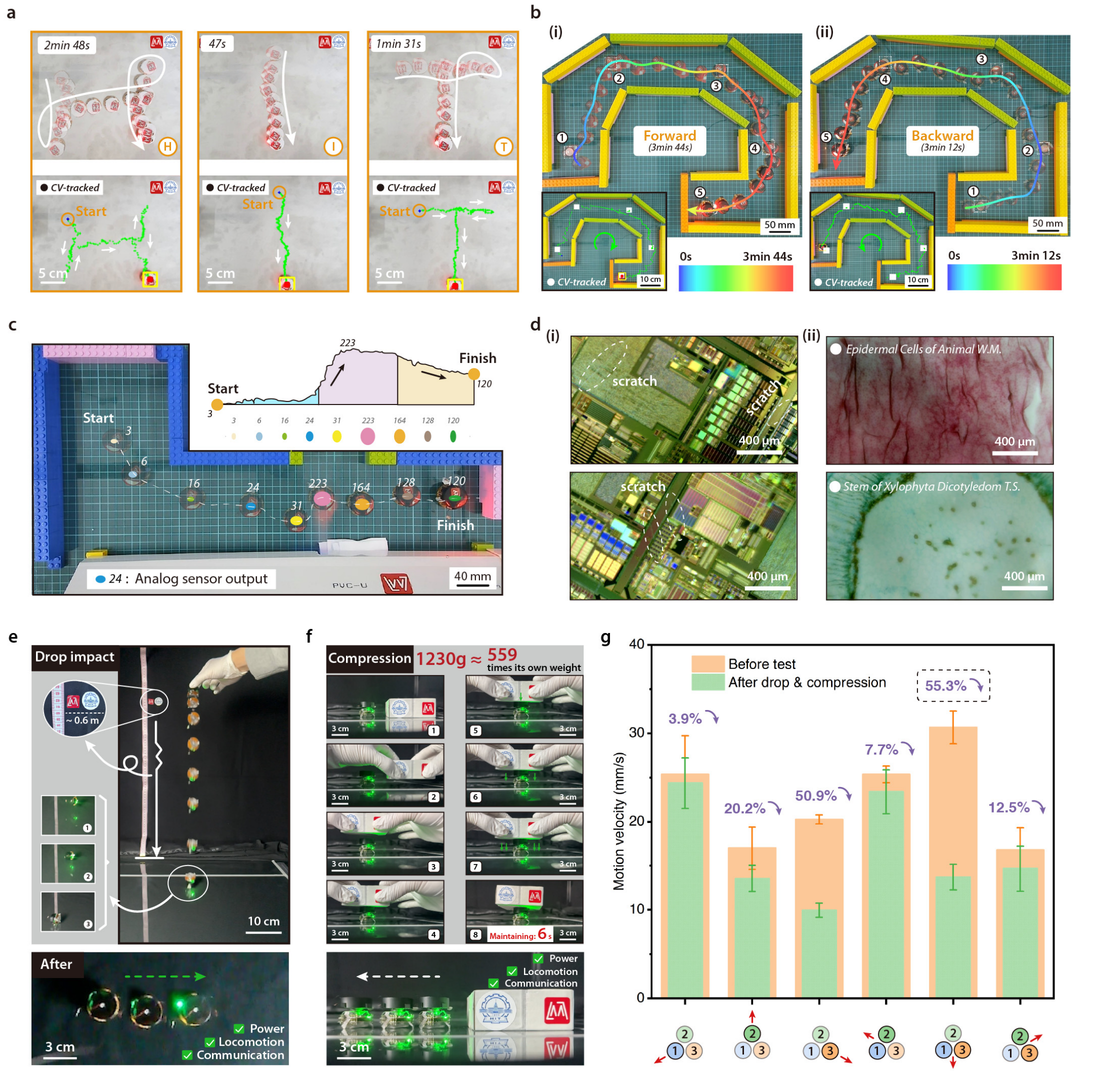
"The robot can carry a maximum payload of approximately 13.63 times its body weight and achieves effective omnidirectional locomotion on glass, marble, and stainless-steel surfaces. Furthermore, the robot retains its functional locomotion capabilities after sustaining a 0.6 m drop impact and a continuous compressive load of approximately 559 times its body weight, demonstrating exceptional structural robustness."

(please see page 20, line 493 of the revised manuscript, marked in yellow)

"iii) Since the robot is driven by micro-amplitude vibrations of its foot tips, its operation has primarily been demonstrated on smooth, planar surfaces. Exploring and expanding the robot's locomotion capabilities on non-uniform surfaces is essential for broadening its potential application scenarios."
(please see page 22, line 522 of the revised manuscript, marked in yellow)



*"**Supplementary Fig. 11.***
***Performance of the robot on diverse contact substrates. (a)** Photographs of the robot situated on glass, marble, and stainless-steel surfaces. **(b)** The velocity characteristics across six motion directions on the three respective substrates. The robot achieves a peak velocity of 30.66 mm/s on the glass substrate (actuation: #1 and #3; LBS: #2). In comparison, the maximum locomotion speeds on marble and stainless-steel surfaces reach 26.21 mm/s and 26.06 mm/s, respectively (actuation: #1 and #2; LBS: #3 for both cases). Note: Data obtained from a representative prototype (Prototype #2, 60 V_{p-p})."*
(please see page 28 of the Supplementary Materials, Supplementary Fig. 11, marked in yellow)



*“**Fig 4. Performance and applications for the robot. (a)** The untethered robot draws the trajectory of “H”, “I”, and “T” letters on the iron substrate. **(b)** The robot demonstrates bidirectional maze navigation and reports five magnetic checkpoints it passed using the onboard hall sensor. (i) Forward. (ii) Backward. **(c)** The gas leak inspection application. (Note: the values only represent the sensor analog values and are not actual gas concentrations) **(d)** The robot is capable of performing tasks in microscopic observation. (i) Wafer detection. (ii) Cell specimen observation (Epidermal Cells of Animal W.M. and Stem of Xylophyta Dicotyledon T.S.). **(e)** The robot (Prototype #2) remains functional after a free-fall impact from a height of 0.6 m onto a desktop. **(f)** The robot (Prototype #2) maintains normal locomotion after sustaining a compressive load of 1230 g (approximately 559 times its body weight). **(g)** Evaluation of velocity reduction after a drop impact and compressive loading (Prototype #2, 60 $V_{p\text{-}p}$; maximum reduction: 55.3% reduction; Actuation: #1 and #3, LBS: #2).”*

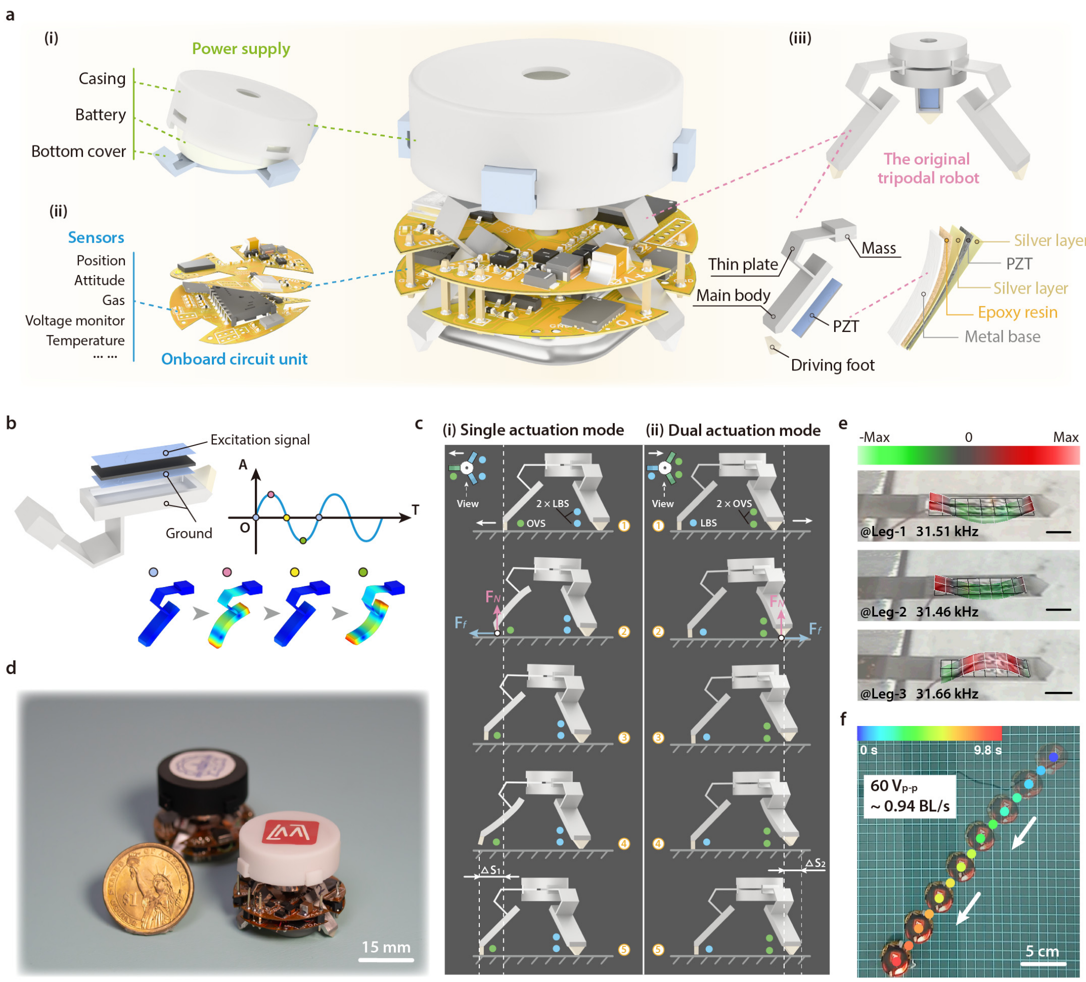
(please see page 37, line 777 of the revised manuscript, Fig. 4, marked in yellow)

Comment 4. Figure 1(c-ii): The two legs are difficult to distinguish. Consider improving contrast or adjusting the view to make the leg structures clearer.

Response: We thank you for this valuable comment.

As you correctly pointed out, due to the 120° spatial distribution of the driving legs and the synchronized movement of the two legs in dual actuation mode, these legs inevitably overlap in the side view. We explored alternative viewing angles; however, none could illustrate the locomotion strategies as intuitively and concisely as the side view. To address the distinguishability issue, we have added an inset top view in the upper-left corner to depict the spatial arrangement and explicitly labeled the motion status (OVS or LBS) of the three legs. Furthermore, we used blue and green markers in the side view to clearly differentiate the status of legs. We have also added a corresponding explanation in the main text to clarify this arrangement.

Modification: The revised **Fig. 1** and its supplementary descriptions have been added to the main text.



"Fig 1. Structure, actuation and locomotion strategies of the robot. ***(a)*** *The structure of the robot and its components. (i) The power supply unit, including a casing, a battery and a bottom cover. (ii) The onboard circuit unit includes wireless drive circuits and five types of sensors. (iii) The tripodal robot.* ***(b)*** *The deformation process of the actuation unit within one exciting signal cycle. (c) Locomotion strategies. (i) Single actuation mode. (ii) Dual actuation mode.* ***(d)*** *A photo of two untethered robots and one dollar.* ***(e)*** *The vibration modes of three driving legs (leg-1: 31.51 kHz, leg-2: 31.46 kHz, leg-3: 31.66 kHz).*

Scale bar: 5 mm. ***(f)*** *Series of optical images recording the untethered motion of a prototype robot in dual actuation mode (60 $V_{p\text{-}p}$, 0.94 BL/s).”*

(please see page 32, line 734 of the revised manuscript, Fig. 1, marked in yellow)

“The states of the driving legs are indicated by green (OVS) and blue (LBS) circles. This figure shows the robot from a lateral view, where the viewing direction is illustrated in the upper-left corner. Under this view, two driving legs are overlapped, and their states are represented by two stacked circles.”

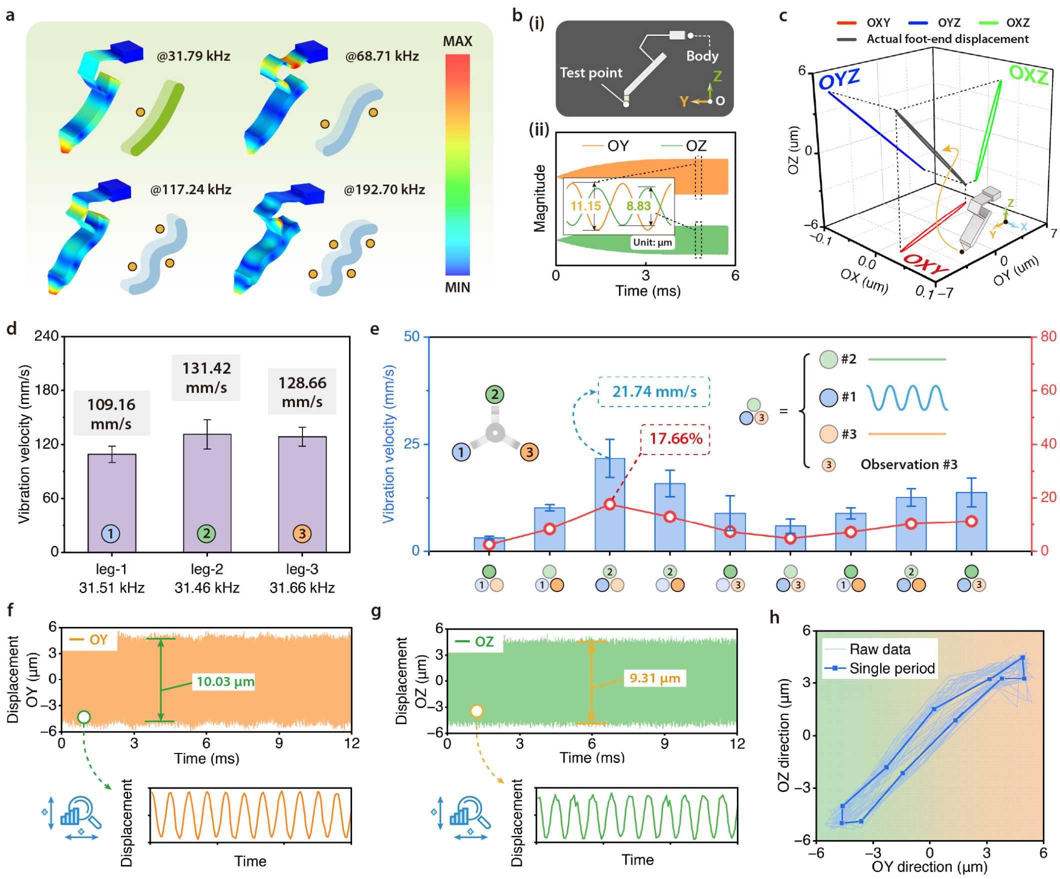
(please see page 6, line 133 of the revised manuscript, marked in yellow)

Comment 5. Figure 2(b): The meaning of the measurements is unclear. Please label the reference frame (O–XYZ) relative to the leg to aid interpretation. It would also help to visually indicate the vibration pattern, rather than showing only a shaded region.

Response: We thank you for this valuable comment.

We have added a schematic diagram of the driving leg and indicated the position of the test point as well as the orientation of the coordinate system. In addition, a locally enlarged view has been included to resolve data overlap caused by too many simulation cycles.

Modification: We have revised the **Fig. 2b** and added the corresponding descriptions in the revised manuscript.



*“****Fig 2. Robot simulation analysis and vibration characteristics.*** ***(a)*** *Bending modes (1 ~ 4) and resonant frequencies of the actuation unit.* ***(b)*** *Transient dynamics simulation results of the robot's foot end. (i) The measurement point is at the foot end. Note: the other two driving legs are hidden for clarity but are*

included in the simulation. (ii) Simulation results: OY: 11.15 μm, OZ: 8.83 μm, 80 V_{p-p}. (c) Spatial trajectory of the robot's foot end. ***(d)*** *Vibration velocity characteristics of driving legs (10 V_{p-p}, leg-1: 109.16 mm/s, leg-2: 131.42 mm/s, leg-3: 128.66 mm/s, fixed frequency mode). Error bars: three experimental trials.* ***(e)*** *The vibration isolation characteristic test (10 V_{p-p}, maximum vibration velocity: 21.74 mm/s, corresponding transmissibility: 17.66%, actuation: #1, observation: #2, fixed frequency mode). Colored circles with 100% and 50% transparency represent the actuation unit under OVS and the non-actuated units, respectively. The numbers on the colored circles indicate the observed actuation unit. Note: the excitation frequency is taken as the mean of their respective resonant frequencies when the two actuation units are required to operate simultaneously.* ***(f)*** *and* ***(g)*** *show the vibration displacements of the prototype foot end in the OY and OZ directions, respectively (OY: 10.03 μm, OZ: 9.31 μm, 80 V_{p-p}).* ***(h)*** *The diagonal vibration trajectory of the foot end. The light blue and dark blue curves represent the raw data and the complete trajectory over a single period, respectively."*

(please see page 33, line 744 of the revised manuscript, marked in yellow)

*"**Figure 2b(i)** defines the orientation of the reference coordinate system, and the test point is located at the foot end. Simulation results show that the vibration amplitudes of the driving foot end along the OY and OZ directions are 11.15 μm and 8.83 μm, respectively, as shown in **Fig. 2b(ii)**."*

(please see page 8, line 185 of the revised manuscript, marked in yellow)

Comment 6. Figure 2(c): Clarify that the three curves are projections onto the plane. Adding simple annotations and a schematic of the leg would make this figure more intuitive.

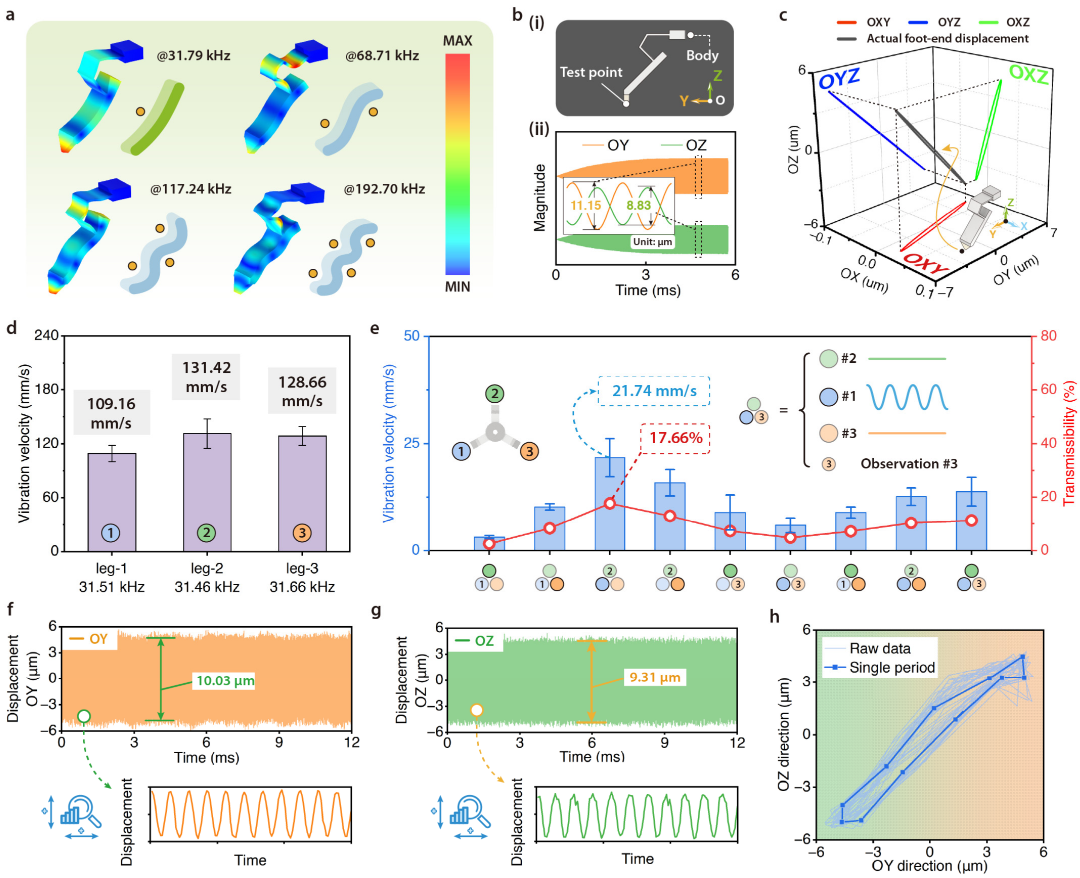
Response: We greatly appreciate your professional comment.

We have updated **Fig. 2c** by adding projection plane labels (e.g., OYZ) and auxiliary projection lines to explicitly illustrate the relationship between the total displacement and the projected displacement. Additionally, we have incorporated data labels, a schematic model of the driving leg, and the coordinate system to enhance clarity. The corresponding descriptions in the main text related to **Fig. 2c** have also been revised and completed accordingly.

Modification: We have revised **Fig. 2c** and added the corresponding explanations to the main text.

*"The vibration trajectory approximates an oblique line, as shown in **Fig. 2c**. The foot-end displacement is mainly concentrated along the OY and OZ directions, while the displacement in the OX direction, which is perpendicular to the driving direction, is extremely small (less than 0.1 μm)."*

(please see page 8, line 188 of the revised manuscript, marked in yellow)



*"**Fig 2. Robot simulation analysis and vibration characteristics.** **(a)** Bending modes (1 ~ 4) and resonant frequencies of the actuation unit. **(b)** Transient dynamics simulation results of the robot's foot end. (i) The measurement point is at the foot end. Note: the other two driving legs are hidden for clarity but are included in the simulation. (ii) Simulation results: OY: 11.15 μm, OZ: 8.83 μm, 80 $V_{p\text{-}p}$. (c) Spatial trajectory of the robot's foot end. **(d)** Vibration velocity characteristics of driving legs (10 $V_{p\text{-}p}$, leg-1: 109.16 mm/s, leg-2: 131.42 mm/s, leg-3: 128.66 mm/s, fixed frequency mode). Error bars: three experimental trials. **(e)** The vibration isolation characteristic test (10 $V_{p\text{-}p}$, maximum vibration velocity: 21.74 mm/s, corresponding transmissibility: 17.66%, actuation: #1, observation: #2, fixed frequency mode). Colored circles with 100% and 50% transparency represent the actuation unit under OVS and the non-actuated units, respectively. The numbers on the colored circles indicate the observed actuation unit. Note: the excitation frequency is taken as the mean of their respective resonant frequencies when the two actuation units are required to operate simultaneously. **(f)** and **(g)** show the vibration displacements of the prototype foot end in the OY and OZ directions, respectively (OY: 10.03 μm, OZ: 9.31 μm, 80 $V_{p\text{-}p}$). **(h)** The diagonal vibration trajectory of the foot end. The light blue and dark blue curves represent the raw data and the complete trajectory over a single period, respectively."*

(please see page 33, line 744 of the revised manuscript, marked in yellow)

Comment 7. Figure 3(a): The subfigures are currently difficult to interpret. The caption mentions "power," but only the top-left subfigure appears to show power data. Please clarify the meaning of the other subfigures and label the vertical axes (especially for the top-right and bottom-right plots).

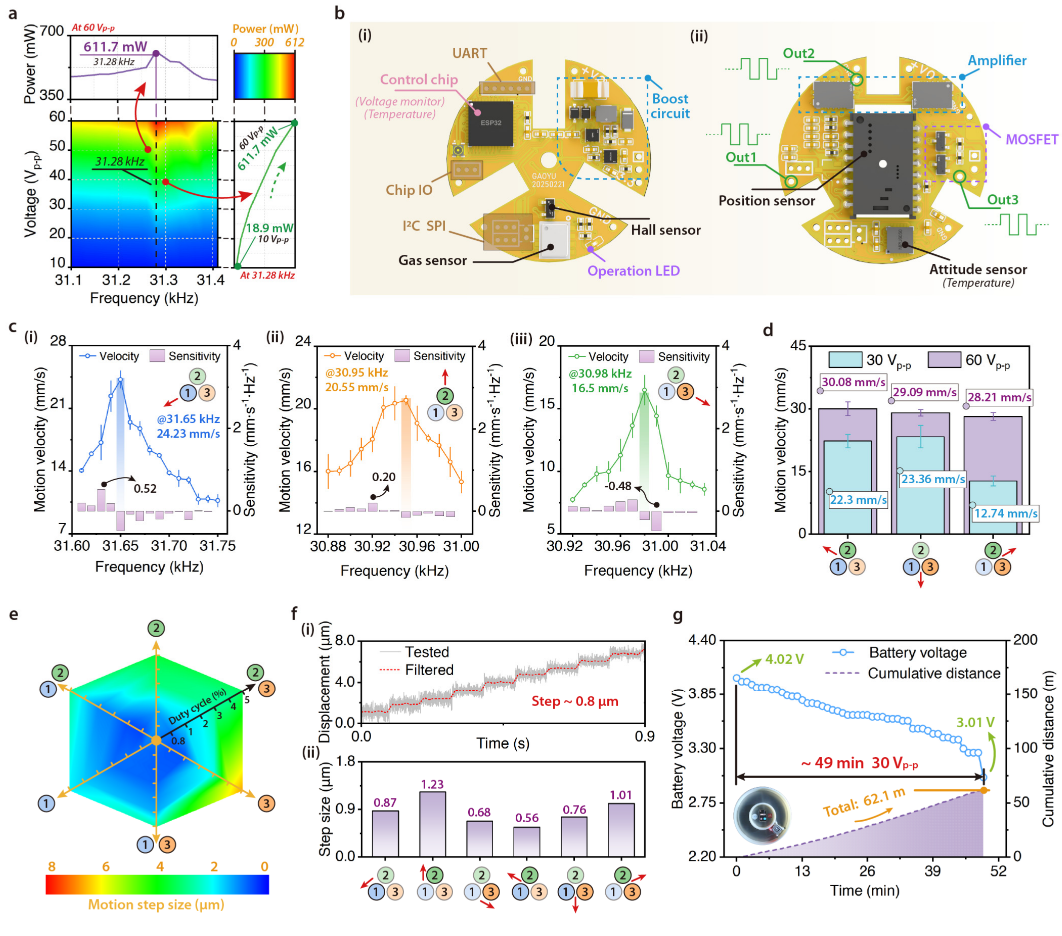
Response: we thank you for this valuable comment.

Fig. 3a consists of four subplots illustrating the power characteristics of the driving leg. Specifically:

- The **top-left** panel shows the relationship between power and frequency at 60 $V_{p\text{-}p}$.
- The **bottom-right** panel displays the relationship between power and excitation voltage at 31.28 kHz.
- The **bottom-left** panel presents the contour plot of power variation with respect to both frequency and voltage.
- The **top-right** panel serves as the corresponding color scale bar.

Due to the same coordinate axis scale, we initially omitted the scale values of the subfigures, which may have caused unnecessary confusion. In the revised **Fig. 3a**, we used dashed lines to clarify the correspondence of the axis scales and labeled the meanings and units of the data axes. We have annotated key data points to enhance readability; for instance, the power values at 10 $V_{p\text{-}p}$ and 60 $V_{p\text{-}p}$ are marked as 18.87 mW and 611.7 mW, respectively. In addition, we have provided a detailed explanation of the meanings of all four subplots in the caption of **Fig. 3a**.

Modification: We have revised **Fig. 3a** and added the corresponding explanations to the main text.



*"**Fig 3. Integration design and locomotion characteristics. (a)** Power consumptions of the tethered robot (single actuation unit, maximum value: 31.28 kHz, 611.7 mW, 60 $V_{p\text{-}p}$). Note: the panels show*

the frequency response at 60 V_{p-p} (top left), the voltage response at 31.28 kHz (bottom right), and the overall power contour map (bottom left) with its corresponding color scale (top right). ***(b)*** *Layout and design of the onboard circuit unit. (i) Top board. (ii) Bottom board.* ***(c)*** *The relationship between velocity and frequency for the untethered robot in single actuation mode (30 V_{p-p}). (i) Actuation: #1, LBS: #2 and #3. (ii) Actuation: #2, LBS: #1 and #3. (iii) Actuation: #3, LBS: #1 and #2. (Error bars: three repeated motion trials.)* ***(d)*** *The relationship between velocity and frequency in dual actuation mode. (Error bars: three repeated motion trials.)* ***(e)*** *The relationship between motion step size and duty cycle in the six directions.* ***(f)*** *Minimum step size in stepping motion. (i) An example displacement data. (ii) Minimum microscopic motion step size in six directions (duty cycle: 0.8%, minimum step size: 0.56 μm). (g) The robot exhibits long endurance over 49 minutes, covering a cumulative travel distance of 62.1 m (initial voltage: 4.02 V, 180 mAh, 30 V_{p-p}).”*
(please see page 35, line 762 of the revised manuscript, marked in yellow)

“*The results demonstrate that power consumption increases with the applied voltage, rising from only 18.87 mW at 10 V_{p-p} to a maximum of 611.7 mW at 60 V_{p-p} (at 31.28 kHz). Under a constant excitation voltage, the power consumption of the driving leg initially increases and subsequently decreases with rising frequency, peaking at the resonance point.”*
(please see page 9, line 214 of the revised manuscript, marked in yellow)

Comment 8. Application Justification: In the section discussing the robot’s use under a microscope for scratch observation, it is unclear why a robot is needed if the microscope can directly observe the scratch. Please clarify the specific advantage or necessity of using the robot in this context.
Response: We sincerely thank you for this valuable comment.
The necessity and advantages of utilizing our robot for microscopic observation are primarily reflected in three aspects: high-resolution motion, omnidirectional mobility without needing to adjust its orientation, and cross-scale workstation switching.

First, the experiment involving scratch detection on a wafer and biological slides microscopic observation (**Fig. 4d** and **Supplementary Movie S8**) is specifically designed to validate the robot's low-speed, high-resolution motion characteristics. While a microscope provides the imaging, the robot acts as a high-precision mobile stage. Leveraging its omnidirectional planar locomotion without the need for orientation adjustment, the robot can reposition itself for observations in different directions seamlessly. This eliminates issues commonly associated with conventional robots, such as blind spots caused by random turning radii and the need for re-localization.

Beyond single-station observation, the robot’s unique ability to switch between microscopic high-precision stepping and macroscopic fast-motion modes enables sophisticated cross-scale microscopic observation (**Supplementary Movie S9**). As shown in **Supplementary Fig. 16a**, a dual-station observation platform consisting of two microscopes with different magnifications (750× and 1500×) is established for image stitching and defect localization. First, the robot, carrying a target chip, moves rapidly to the scanning station (750×). Due to the limited field of view (FOV) inherent to high-magnification lenses, the robot utilizes its high-precision stepping mode combined with an image-stitching algorithm to reconstruct a panoramic view of the chip (**Supplementary Fig. 16b**). The stitched results reveal two potential defects consisting of suspected broken wires in this wire-bonding chip.

Subsequently, the robot switches to macroscopic fast-motion mode to reach the second workstation—the inspection station (1500×). Here, the robot reverts to microscopic stepping mode to precisely relocate and confirm the identified defects. Using the stitched panoramic image as a reference for planar coordinates, the robot successfully confirms two wire-bonding fractures (**Supplementary Fig. 16c**). This experiment demonstrates the robot's flexible cross-scale motion switching capability and its promising potential in fields such as micro-defect inspection.

Looking forward, our research is no longer limited to passive observation. We are actively working toward autonomous micro-operation. By integrating the robot with active micro-grippers, we aim to demonstrate more complex macro-micro cross-scale applications, such as micro-structure docking and microsphere handling. Furthermore, we are exploring multi-robot collaborative scenarios for coordinated task execution, which we believe will expand the practical utility of this platform.

Modification: We have added the relevant description to the revised manuscript.

*"In addition to high-resolution microscopic observation, the robot can leverage its ability to switch between microscopic high-precision and macroscopic fast-motion modes to perform cross-scale microscopic observation (see **Supplementary Movie S9**). As shown in **Supplementary Fig. 16a**, a dual-station observation platform consisting of two microscopes with different magnifications (750× and 1500×) is established for image stitching and defect localization. First, the robot, carrying a target chip, moves rapidly to the scanning station (750×). Due to the limited field of view (FOV) inherent to high-magnification lenses, the robot utilizes its high-precision stepping mode combined with an image-stitching algorithm to reconstruct a panoramic view of the chip (**Supplementary Fig. 16b**). The stitched results reveal two potential defects consisting of suspected broken wires in this multi-pin wire-bonding chip.*

*Subsequently, the robot switches to macroscopic fast-motion mode to reach the second workstation—the inspection station (1500×). Here, the robot reverts to microscopic stepping mode to precisely relocate and confirm the identified defects. Using the stitched panoramic image as a reference for planar coordinates, the robot successfully confirms two wire-bonding fractures (**Supplementary Fig. 16c**). This experiment demonstrates the robot's flexible cross-scale motion switching capability and its promising potential in fields such as micro-defect inspection."*

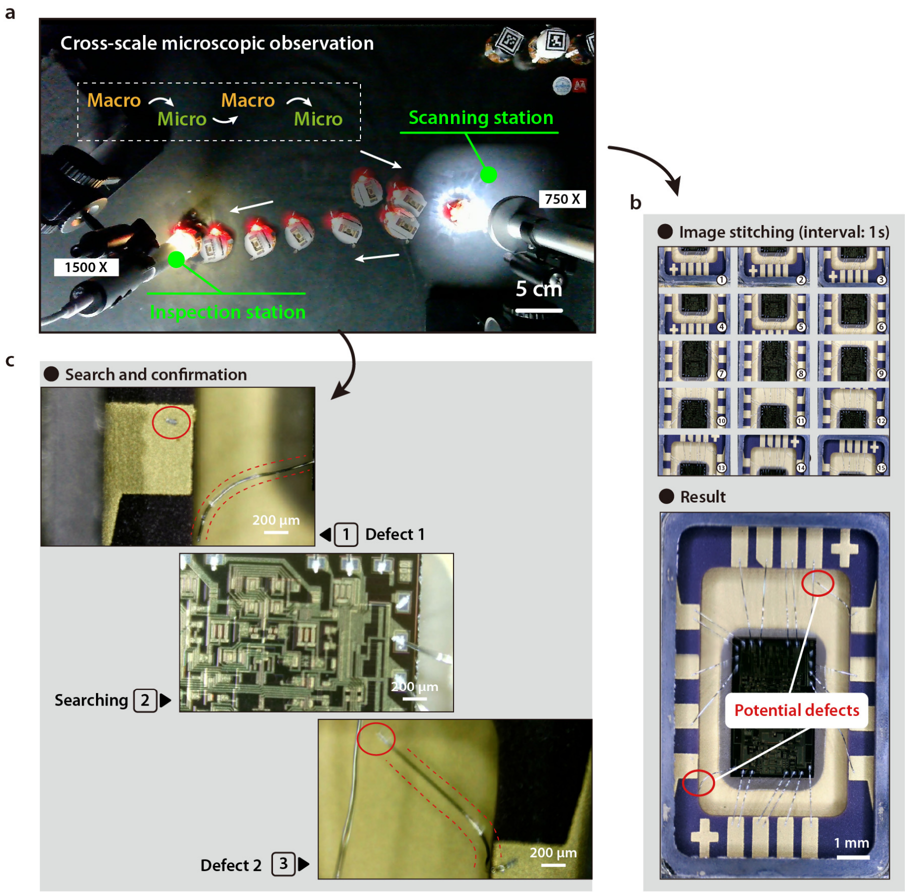
(please see page 15, line 364 of the revised manuscript, marked in yellow)

"Moreover, the robot can switch between a continuous mode with speeds up to 3 cm/s for maze navigation, and an intermittent mode with a minimum step size of 0.56 μm for precision microscopic observation. This cross-scale locomotion capability enables the robot to efficiently perform wide-range, multi-point chip defect inspections."

(please see page 20, line 486 of the revised manuscript, marked in yellow)

"ii) Integrating functional operational units (such as miniature grippers) to enable micro-gripping applications and multi-robot collaborative demonstrations."

(please see page 22, line 531 of the revised manuscript, marked in yellow)



*"**Supplementary Fig. 16.***

***Cross-scale microscopic observation application. (a)** Global camera view for cross-scale operation. The robot moves to the scanning station, where it utilizes its high-precision stepping mode to perform image stitching of a target chip. Then it switches to fast macroscopic locomotion to reach the inspection station for detailed defect localization. **(b)** Raw sequential frames (interval: 1 s) and the stitched result. Two potential defects (Suspected breakage and detachment of bonding wire) are identified within the stitched panoramic image. **(c)** Defect search and identification. Detailed localization and confirmation of the two identified defects are performed at the inspection station."*

(please see page 33 of the Supplementary Materials, Supplementary Fig. 16, marked in yellow)

Responses to reviewers' comments (Reviewer: #1)

Remarks to the Author:

I thank the authors for the revised manuscript and my concerns have been addressed in full. I recommend publication of this work.

Response: We sincerely thank the reviewer for evaluating our revised manuscript and for the positive feedback. We deeply appreciate the time and effort you have dedicated to reviewing our work and for recommending it for publication. Your constructive comments during the revision process have been invaluable.

Responses to reviewers' comments (Reviewer: #2)

Remarks to the Author:

The reviewer appreciates the authors' substantial efforts in addressing the previous comments. All concerns have been satisfactorily resolved, and the reviewer has no further suggestions for improvement.

Response: We are very grateful to the reviewer for recognizing our efforts during the revision process. We sincerely appreciate your thoughtful and constructive feedback, which has significantly helped us improve the overall quality and rigor of our manuscript.