

Tunable Impact and Vibration Absorbing Neck for Robust Visual-Inertial State Estimation for Dynamic Legged Robots

Taekyun Kim [✉], Sangbae Kim, *Member, IEEE*, and Dongjun Lee [✉], *Member, IEEE*

Abstract—We propose a new neck design for legged robots to achieve robust visual-inertial state estimation in dynamic locomotion. While visual-inertial state estimation is widely used in robotics, it has a problem of being disturbed by the impacts and vibration generated when legged robots move dynamically. The use of rubber dampers may be a solution, but even if the dampers are proper for some gaits, they may be excessively deformed or resonated at certain frequencies during other gait locomotion since they are not tunable. To address this problem, we develop a tunable neck system that absorbs the impacts and vibration during diverse gait locomotions. This neck system consists of two components: 1) a suspension mechanism that compensates for the weight of the head equipped with a camera and IMU (inertial measurement unit), absorbs the impacts and the head motion of high frequencies including vibration as a fixed low-pass filter; and 2) a dynamic vibration absorber (DVA) that can be reactively-adjusted to diverse gait frequencies to alleviate excessive head movements. We present a dynamics analysis of the neck system and show how to adjust the target frequency of the system. Simulation and experimental validation are performed to verify the effect of the proposed neck design, manifesting superior estimation performance and robustness across diverse gaits.

Index Terms—Legged Robots, Mechanism Design, Visual-Inertial SLAM.

I. INTRODUCTION

LEGGED robots have been widely researched due to their capability to operate in irregular terrain. One of the central issues in legged robots is precise state estimation for their safe autonomy. In general, the state estimation using cameras and IMUs (inertial measurement units) is widely exploited due to their low cost and complementary characteristics (e.g., indoor

UAV [1], [2], hand motion tracking [3]). Likewise, for the state estimation of legged robots, several studies have developed methods using visual-inertial information with leg kinematic information [4], [5], [6], which circumvent the drift issue of methods using only proprioceptive sensors (i.e. IMU and encoder) [7], [8], [9].

These state estimation methods using visual-inertial measurements, however, are fragile during aggressive dynamic locomotions such as pronking gait. Specifically, if the impact is large and/or generates high-frequency vibrations, the accelerometer signal is in general compromised due to saturation and the limited sampling rate. For drones, which are also subject to the vibrations caused by propellers, rubber dampers are typically used to absorb the vibration as a mechanical low-pass filter. However, in the case of legged robots, even if certain rubber dampers are appropriate for some gaits, they may be resonated at gait locomotion of other frequencies since they are not tunable. In addition, while low-stiffness dampers are more advantageous in absorbing vibrations, some large motions would deform excessively these dampers. As another way to improve the robustness of state estimation, some legged robots adopt multiple cameras and/or LiDAR (light detection and ranging) sensor (e.g., a camera with Velodyne LiDAR sensor for Anymal [10], five cameras for Spot [11]), which result in high hardware and system integration costs.

To reduce impacts or vibrations induced by aggressive motions of legged robots, some studies have focused on inserting compliant components in the legs (e.g., flexible feet [14], elastic actuators [15], [16]). However, these components are not tunable once installed, so the stiffness of the components may be effective in improving state estimation for some locomotions but not for others. Another recent work [17] presents a shock propagation reduction method by optimizing the mass distribution in a robot's leg, which, however, cannot be applied to various gaits since it restricts the ankle angle on landing. On the other hand, an algorithmic method is proposed in [18], which enhances the state estimation during aggressive locomotion by performing multiple visual SLAM sessions that track feature points individually according to the phases of the robot's gait cycle. This, however, would be less accurate and robust than using the continuous image frames as enabled by our proposed neck mechanism since only IMU measurements are used to constrain separate visual SLAM sessions.

Manuscript received 26 September 2022; accepted 21 January 2023. Date of publication 27 January 2023; date of current version 2 February 2023. This letter was recommended for publication by Associate Editor K. Hashimoto and Editor A. Kheddar upon evaluation of the reviewers' comments. This work was supported in part by the Korea Institute for Advancement of Technology funded by the Ministry of Trade, Industry & Energy (MOTIE) of Korea, through the HRD Program for Industrial Innovation under Grant P0020536, and in part by the National Research Foundation (NRF) funded by the Ministry of Science and ICT (MSIT) of Korea, through the Basic Science Research Program and the Engineering Research Center Program under Grants 2020R1A2C301003913 and 2016R1A5A193847223. (*Corresponding author: Dongjun Lee.*)

Taekyun Kim and Dongjun Lee are with the Department of Mechanical Engineering, IAMD and IER, Seoul National University, Seoul 08826, Republic of Korea (e-mail: ktk4532@snu.ac.kr; djlee@snu.ac.kr).

Sangbae Kim is with the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: sangbae@mit.edu).

Digital Object Identifier 10.1109/LRA.2023.3240369



Fig. 1. A sprinting cheetah. Four snapshots show that the head moves stably even with the dynamic movement of the body. Image courtesy of National Geographic [12].

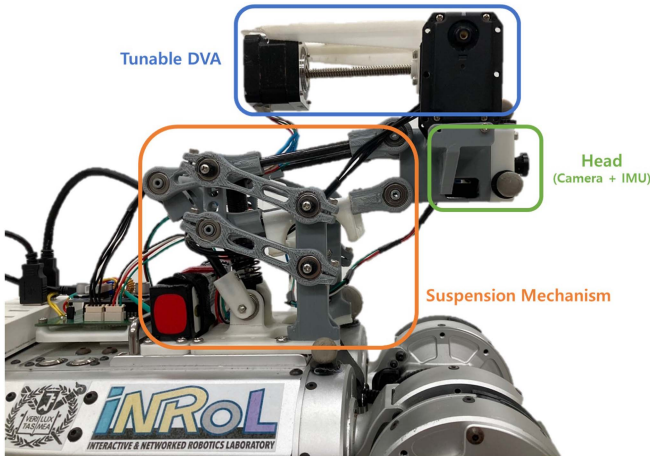


Fig. 2. The prototype of proposed neck design with a Mini Cheetah [13].

In this letter, we pay attention to the neck as a solution to the impact and vibration problem for the state estimation of legged robots. In fact, this approach is inspired by nature, where some animals decouple the movement of the head with a cognitive sensor from that of the body. For example, as shown in Fig 1, a cheetah can keep its head stably when sprinting toward prey. In the same way, for the stable perception of legged robots, we consider the use of the neck desirable, which can decouple the head motion from the body motion.

More specifically, as shown in Fig. 2, we first design a suspension mechanism that compensates for the weight of the head equipped with a camera and IMU. With a spring and a damper, this suspension mechanism absorbs the impacts from the legs and provides a response of a low-pass filter to the head motion. Although this mechanism can absorb the impact and vibration from the legs in some gaits like rubber dampers, it may still be resonated or excessively deformed for other gaits. If the head with the sensors excessively moves, it may collide with the range of motion limit, disturbing the accelerometer signal of the IMU. Also, the camera may suffer from motion blur rendering feature tracking more difficult or impossible. Therefore, we adopt a tunable dynamic vibration absorber (DVA) to stabilize the head motions, which is widely utilized to mechanically stabilize oscillations such as skyscrapers [19] and rotating machinery [20]. To handle diverse locomotion, the DVA can adjust its notching frequency according to the gait frequency, thereby alleviating

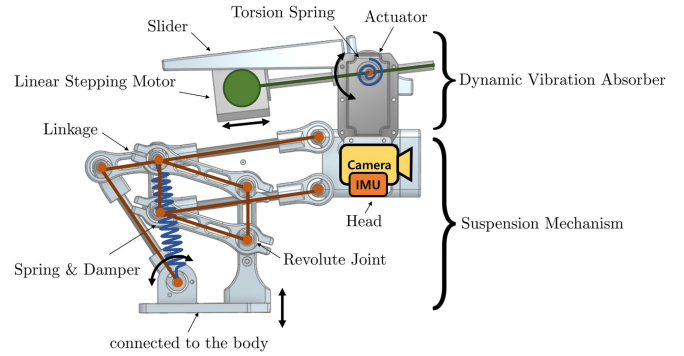


Fig. 3. The overall proposed neck design and its main components.

the excessive head motion. This DVA mechanism turns out to be necessary during some aggressive locomotion in which the head motion can become very excessive if only the suspension is used.

As an application of morphological computation [21], our proposed design replaces the complex computation of active control with passive components for shock and vibration absorption. In addition, these (analog) passive elements provide a much faster response as compared to (digital) active control with a finitely-fast sampling rate. In the same context, in contrast to using some deep learning-based methods for image stabilization (e.g., DeepDeblur [22]), our approach can mechanically replace their computational time/resources to obtain stable images. Of course, our mechanism is much simpler as compared to the neck of real animals with real-time variable stiffness. Even so, we believe this letter manifests the feasibility of (simpler) DVA-based neck design for better visual-inertial state estimation of legged robots. This perception-stabilizing efficacy of our DVA-based neck design is verified with simulations and experiments using MIT mini cheetah [13] with a prototype of our neck design shown in Fig 2. To the best of our knowledge, this is the first result of exploring the possibility of utilizing a neck mechanism to improve the perception stability and performance for dynamic legged robots.

The rest of the letter is organized as follows. In Section II, the neck design is proposed in detail. Section III and Section IV provide an analysis and hardware implementation of our neck mechanism, respectively. The simulation and experimental results are presented in Section V and Sec VI, followed by some concluding remarks and comments on future works in Section VII.

II. NECK MECHANISM DESIGN

As seen in Fig. 3, the proposed neck design is mainly composed of two parts: 1) the suspension mechanism that compensates for the weight of the head, maintains the head orientation to be the same as that of the body, and absorbs impacts and motions of high frequency including vibrations; and 2) the DVA which can be adjusted to the gait frequencies to suppress excessive motion and vibration of the head with the camera and IMU attached on it.

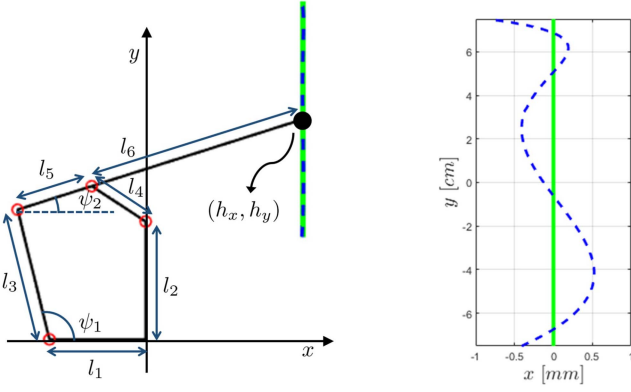


Fig. 4. (left) Kinematics model and design parameters of the proposed neck design. (right) The desired vertical trajectory and the trajectory of the head with the optimized design parameters represented with respect to desired trajectory center (x_d, y_{mid}) , where $y_{mid} = \frac{1}{2}(y_{min} + y_{max})$. The green line denotes the desired vertical trajectory of the head represented by (x_d, y_{min}, y_{max}) and the dashed blue line indicates the trajectory of the head determined by the geometric parameters l_i .

A. Suspension Mechanism Design

The main part of the suspension consists of a linkage mechanism connecting the body and head, which imitates the neck joint and skeleton of animals. We use a compression spring and a hydraulic damper to perform the roles of muscles in this mechanism. When absorbing an impact and vibration, this linkage should operate in a vertical direction, which is the main direction of the impact and vibration, not in any other unnecessary direction. For this vertical movement of the head, we generalize the leg design of Ascento [23], which was developed to create a vertical trajectory of the foot passing the robot's center of mass, in a way that places the trajectory forward as much as we want so as not to block the view of the camera, as seen in Fig. 4. This linkage mechanism cannot make the vertical movement of the head perfectly because the bars of the linkage are connected via revolute joints. Thus, we optimize the design parameters of the linkage for the head trajectory to track the desired vertical trajectory as much as possible.

For this, with the geometric design parameters l_i and the joint angles ψ_i in Fig. 4, note that the position of the head (h_x, h_y) is represented as

$$h_x(\psi_1, \psi_2, l_i) := -l_1 + l_3 \cos \psi_1 + (l_5 + l_6) \cos \psi_2 \quad (1)$$

$$h_y(\psi_1, \psi_2, l_i) := l_3 \sin \psi_1 + (l_5 + l_6) \sin \psi_2 \quad (2)$$

and the constraint of the kinematic loop can be expressed as

$$c_{loop}(\psi_1, \psi_2, l_i) = 0 \quad (3)$$

where $c_{loop}(\psi_1, \psi_2, l_i)$ is defined by

$$c_{loop}(\psi_1, \psi_2, l_i) := (l_3 \cos \psi_1 + l_5 \cos \psi_2 - l_1)^2 + (l_3 \sin \psi_1 + l_5 \sin \psi_2 - l_2)^2 - l_4^2.$$

To express the desired vertical trajectory, we use the parameters of the forward distance from the body to the head x_d and the minimum/maximum height of the head y_{min}, y_{max} , as seen in

TABLE I
DESIRED TRAJECTORY PARAMETERS

Parameter	Length [mm]
x_d	48
y_{min}	80
y_{max}	200

TABLE II
DETERMINED DESIGN PARAMETERS FOR PROTOTYPE LINKAGE

Parameter	Length [mm]
l_1	47.0
l_2	63.3
l_3	92.2
l_4	62.6
l_5	38.2
l_6	108.7

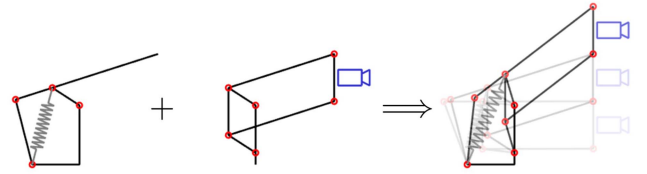


Fig. 5. The diagram of the suspension mechanism. The entire neck structure consists of a combination of an impact-absorbing linkage and two connected parallel four-bar linkage.

Fig. 4. For the prototype design used in experimental validation, we determine these desired trajectory parameters in Table I considering that the neck can cover large motions such as pronk, the head should not collide with other parts of the robot, and the camera's field of view should not be blocked.

With the head trajectory derived in (1)–(3), we can formulate the following design optimization to minimize the deviation from the desired vertical motion of the head:

$$\operatorname{argmin}_{l_i} \int_{y_{min}}^{y_{max}} (x - x_d)^2 dy \quad (4)$$

$$\text{subject to } x = h_x(\psi_1, \psi_2, l_i), \quad (5)$$

$$y = h_y(\psi_1, \psi_2, l_i), \quad (6)$$

$$c_{loop}(\psi_1, \psi_2, l_i) = 0 \quad (7)$$

$$l_i \in [0, l_{max}] \quad (8)$$

where: (5)–(7) is to constrain the head position to be in the trajectory generated by l_i ; and (8) is to ensure the positive length of the bars and to prevent this optimization from diverging to infinite length. Through the design optimization (4)–(8), the geometric parameters are determined in Table II for the desired vertical trajectory given in Table I. With these design parameters, the deviation of the head is horizontally bounded in 1 mm when it moves 12 cm along the vertical direction, as shown in Fig. 4.

Moreover, we supplement this impact-absorbing linkage with the two connected parallel four-bar linkage, see Fig. 5. Due to this additional parallel linkage, the rotation of the neck joints does not change the direction of the head. In other words, the head and the body always have the same orientation.

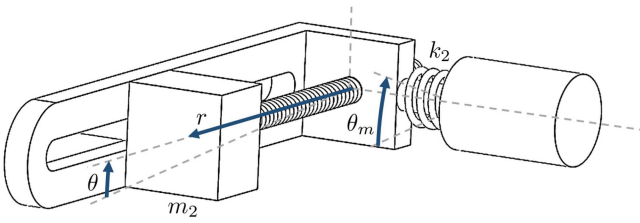


Fig. 6. The diagram of tunable dynamic vibration absorber (DVA). By changing the position (r) of the linear stepping motor along with the slider, it can adjust the natural frequency.

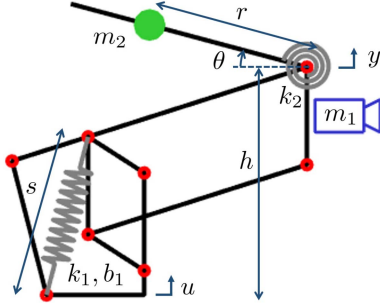


Fig. 7. The diagram of dynamics model for the neck design.

Here, note that (3) defines a constraint for ψ_1 and ψ_2 . This implies that only one of either ψ_1 or ψ_2 is independent. Further, we only need just one rotary encoder to calculate the relative position of the head from the body (i.e., (1)–(2)), and we already know that the head and the body have the same orientation. Using this kinematic information between the head and the body, the estimated state from the sensors on the head can be exploited for the control and planning of the robots.

B. Tunable Dynamic Vibration Absorber

The proposed DVA consists of an actuator, a linear stepping motor, a torsion spring, and a slider. A torsion spring connects the actuator to the slider, similar to series elastic actuators [24], as shown in Fig. 6. By changing the position of the linear stepping motor along with the slider, the DVA can adjust the notch frequency of the system. Then, the actuator rotates the slider to maintain it horizontally in a steady state. The analysis of this tunable DVA is presented in Section III in detail.

III. NECK DYNAMICS ANALYSIS

This section provides the dynamics model to verify that the proposed neck system renders the mechanical notch filter. Furthermore, the method of matching target frequencies is also presented.

We can represent the whole neck system as a mass-spring-damper shown in Fig. 7, viewing the height of the body u as the input and the height of the head y as the output of the system.

We define the generalized coordinates as $q = [y \ \theta]^T$, where θ denotes the rotated angle of the slider. The angle of the actuator is expressed in θ_m (see Fig. 6), and r implies the displacement of

the linear stepping motor of DVA. The mass of the head and the linear stepping motor of DVA are indicated by m_1, m_2 modeled as entirely located at the center of mass, and the other masses are neglected. The spring and damping coefficients of the shock absorber are represented in k_1 and b_1 and the spring coefficient of the DVA is implied by k_2 . The relative position of the head from the body in a vertical direction can be expressed as

$$h(y, u) = h_0 + y - u \quad (9)$$

where h_0 implies the value of $h(y, u)$ in steady state ($y, u = 0$). With the assumption that the horizontal motion of the head is negligible as shown in Fig. 4, the spring/damper length of suspension can be represented by h as

$$s(h) = \sqrt{s_0^2 + \frac{l_5}{l_5 + l_6}(h^2 - h_0^2)} \quad (10)$$

where s_0 denotes the value of $s(h)$ in steady state ($h = h_0$).

The kinetic energy and potential energy of the system can be derived from the states and parameters described above, which lead to the dynamics equation of the system by using the Euler-Lagrange equation

$$\begin{aligned} (m_1 + m_2)\ddot{y} + m_2 r \cos \theta \ddot{\theta} - m_2 r \sin \theta \dot{\theta}^2 \\ + (m_1 + m_2)g + k_1(s - \bar{s}) \frac{\partial s}{\partial y} = \tau_y \\ m_2 r \cos \theta \ddot{y} + m_2 r^2 \ddot{\theta} + k_2(\theta - \theta_m) + m_2 r g \cos \theta = 0 \end{aligned} \quad (11)$$

where \bar{s} is the free length of the spring, and τ_y is the generalized force of the damper associated with the virtual displacement δy . From the virtual work of the damper δW , we can find τ_y as follows:

$$\begin{aligned} \delta W &= -b_1 \dot{s} \delta s \\ &= \underbrace{-b_1 \dot{s} \frac{\partial s}{\partial y}}_{=: \tau_y} \delta y - b_1 \dot{s} \frac{\partial s}{\partial u} \delta u \end{aligned} \quad (12)$$

On the other hand, using the vertical deviation of the head from the steady state ($\Delta h := h - h_0 = y - u$), (10) can be expressed as

$$\frac{\Delta s}{\Delta h} = \frac{2h_0 + \Delta h}{2s_0 + \Delta s} \cdot \frac{l_5}{l_5 + l_6} \quad (13)$$

with $\Delta s := s - s_0$. In the neighborhood of the steady state ($|\Delta s| \ll s_0, |\Delta h| \ll h_0$), we can approximate (13) showing the linear relation between $\Delta s, \Delta h$ as

$$\frac{\Delta s}{\Delta h} \approx \underbrace{\frac{h_0}{s_0} \cdot \frac{l_5}{l_5 + l_6}}_{=: \alpha} \quad (14)$$

Using this approximate relation, the first line of (11) can be approximated as

$$\begin{aligned} (m_1 + m_2)\ddot{y} + m_2 r \cos \theta \ddot{\theta} - m_2 r \sin \theta \dot{\theta}^2 \\ + (m_1 + m_2)g + k_1 \alpha (s - \bar{s}) = -b_1 \alpha^2 (\dot{y} - \dot{u}) \end{aligned} \quad (15)$$

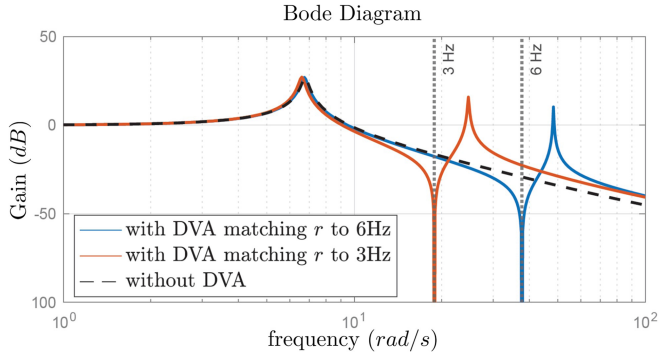


Fig. 8. The bode plots of linearized neck system with/without a DVA, generated by the transfer function of (19).

With the steady-state condition at $\theta = 0$, $s = s_0$, we can derive equations representing relationships as:

$$(m_1 + m_2)g + k_1\alpha(s_0 - \bar{s}) = 0 \quad (16)$$

$$m_2rg - k_2\theta_m = 0 \quad (17)$$

In (17), we can find the motor angle θ_m to make a steady state at $\theta = 0$. In addition to (16), (17), by assuming the small rotated angle/velocity ($\theta, \dot{\theta}$), the dynamics equation (15) can be linearized as

$$\begin{aligned} (m_1 + m_2)\ddot{y} + m_2r\ddot{\theta} + k_1\alpha^2(y - u) &= -b_1\alpha^2(\dot{y} - \dot{u}) \\ m_2r\ddot{y} + m_2r^2\ddot{\theta} + k_2\theta &= 0. \end{aligned} \quad (18)$$

Through this linearized dynamics equation (18), we can derive the transfer function of the system as follows:

$$H(s) = \frac{(m_2r^2s^2 + k_2)(b_1s + k_1)\alpha^2}{D(s)},$$

$$\text{where } D(s) = D_1(s)D_2(s) - m_2^2r^2s^4,$$

$$D_1(s) = (m_1 + m_2)s^2 + b_1\alpha^2s + k_1\alpha^2,$$

$$D_2(s) = m_2r^2s^2 + k_2. \quad (19)$$

The transfer function $H(s)$ has the zero of $w = \sqrt{\frac{k_2}{m_2r^2}}$, which implies the notching frequency. That is, for the target frequency of the body vertical motion w_t , we can set the $r = \sqrt{\frac{k_2}{m_2w_t^2}}$ so as to absorb the oscillation of the head. Fig. 8 shows bode plots according to the step motor position adjusted to the target frequencies.

IV. HARDWARE IMPLEMENTATION

For the suspension mechanism, we use the small shock absorber made for the same purpose in an RC vehicle. It can compress the spring by tightening the nut, allowing it to adjust the height of the head in a steady state. As the suspension system drives the movement of the head like a second-order low pass filter for body movement, we choose a spring coefficient as low as possible under the condition of lifting the head to lower the cutoff frequency. For the actuator of the DVA, we used a Robotis'

TABLE III
INPUT MODELS AS VERTICAL MOTION IN SIMULATION

Gait Type	Frequency [Hz]	Amplitude [mm]
Trot	6	0.15
Pronk	3	0.25

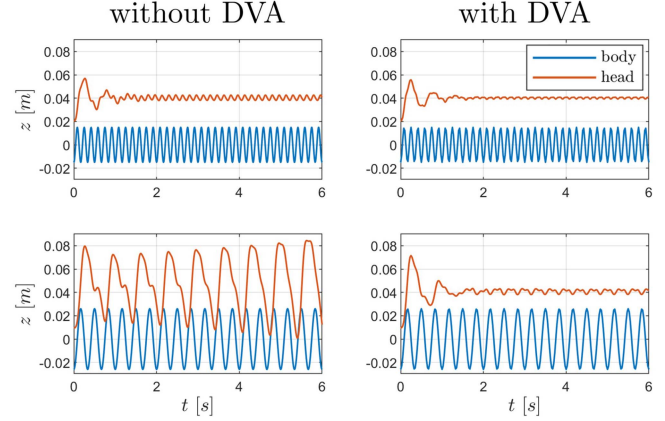


Fig. 9. The z position of body and head in the simulation (left) without DVA and (right) with DVA. The simulation is conducted in two input models: (top) trot and (bottom) pronk.

Dynamixel MX-64 that could withstand the torques caused by the rotating part.

The proposed neck platform should be as light as possible to minimize the effect on the control of the robot. It is crucial for the assumption that the legged robot is not affected by the neck dynamics. At the same time, the neck requires appropriate stiffness of the structure to resist the deformation caused by impact and movement. Therefore, the main simple parts consist of lightweight and high-stiffness carbon fiber tubes. In addition, 3D printing with ABS filament is primarily employed for the geometrically complex connection parts. Under these design considerations, we developed the prototype of the proposed neck design for MIT mini Cheetah [13], as seen in Fig. 2, which results in the overall weight of 0.92 kg including DVA (9% of the robot).

V. SIMULATION

To verify our proposed method, we conduct a simulation showing the behavior of the head, depending on whether or not using the DVA adapted to the target frequency. By numerical integration of the original dynamics equation, we simulate the head motion according to the given vertical movement of the body with MATLAB. In this simulation, we use the same parameters in Table II and two sinusoid input models similar to the data of real robot motion, seen in Table III.

Fig. 9 shows the simulation results about the head motion for the body motion. For the trot input model, the head motions are more smooth than the body in both cases. This is because the trot input has a high frequency and small amplitude, so the suspension mechanism can alleviate the periodic motion of the head, which can be predicted from the bode plot in Fig. 8. On the other side, in the case of the pronking gait, the head motions become unstable without the DVA, which might be caused by the

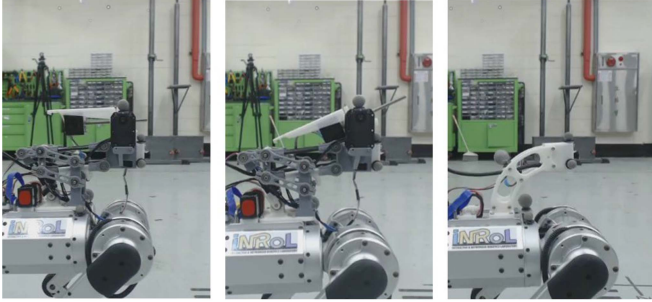


Fig. 10. Three necks used in experiments: (left) **DVA**, (middle) **SUS**, and (right) **FIX**.

resonance of the system or the nonlinear terms neglected when simplifying the dynamics equation. In contrast, we confirm that the DVA can render stable head movement because it absorbs the oscillation even with large body motions of low frequency.

VI. EXPERIMENTS

In this section, we present experimental results with the proposed prototype of the neck design. Each experiment is conducted with an MIT mini Cheetah [13] which has strength in dynamic locomotion. We manually¹ control our robot based on [25], which effectively combines model predictive control (MPC) and whole-body impulse control (WBIC). This controller generates a swing foot trajectory between the footstep locations based on Bezier curves. We use two types of gait in the experiments, specified by two parameters described in [26]; phase offset and stance period for each foot.

- *Trot*: The diagonally symmetrical two legs have the same phase offset. Legs in different groups have a half-phase difference from each other.
- *Pronk*: All legs work in the same phase of the cycle.

And we set the stance period of both gaits to half of the cycle. The frequency of these two gaits is set to 3 Hz, but the trotting gait has a vertical motion of 6 Hz because it contains two times of ground pushing in a cycle. All sensor data are captured using an MPU6050 IMU and an oCam-1MGN-U global shutter camera.

To compare the effect of the proposed neck mechanism on the state estimation of dynamic legged robots, we use three necks as seen in Fig. 10.

- 1) *DVA*: the prototype of the proposed neck design including the DVA, which is adjusted to the target frequency.
- 2) *SUS*: the prototype of the proposed neck design, but using only the suspension mechanism with the DVA locked.
- 3) *FIX*: the neck is replaced by a rigid fixture with a similar dimension.

A. Stabilization of Sensor Data

One of the goals to use the neck mechanism is to stabilize the periodic vertical movement of legged robots. To validate

¹This manual control generates some deviation in the trajectories the robot follows. However, the performance difference with the neck (Fig. 14, Fig. 15) is too drastic to be typically attributable to such (rather minor) executed trajectory differences.

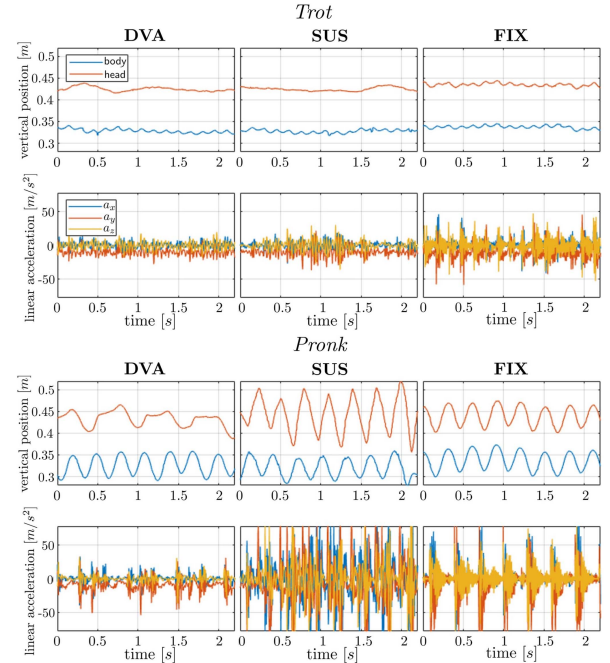


Fig. 11. The vertical position data of the head and linear acceleration data measured by an IMU. These experiments are conducted in two gaits: (top) trot and (bottom) pronk.

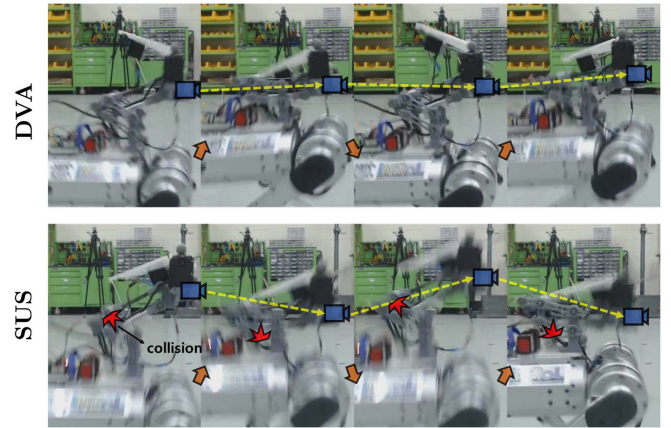


Fig. 12. Snapshots showing the motion of the neck, in the case of pranking gait with (top) **DVA** and (bottom) **SUS**.

the effect of the suspension mechanism and DVA on vertical movements, the OptiTrack MOCAP (motion capture) system measures the motions of the head and body at 100 Hz. Furthermore, the linear acceleration data on the head is measured to see how much shock and vibration it absorbs, which is critical for robust visual-inertial state estimation. We also collect camera image data to see how much excessive motions of the camera affect the performance of feature tracking by comparing how long a thousand feature points of the first image are tracked. The results are shown in Fig. 11 (vertical movement and IMU signal), Fig. 12 (snapshots with **DVA** and **SUS**), and Fig. 13 (feature tracking performance with **DVA**, **SUS**, and **FIX**).

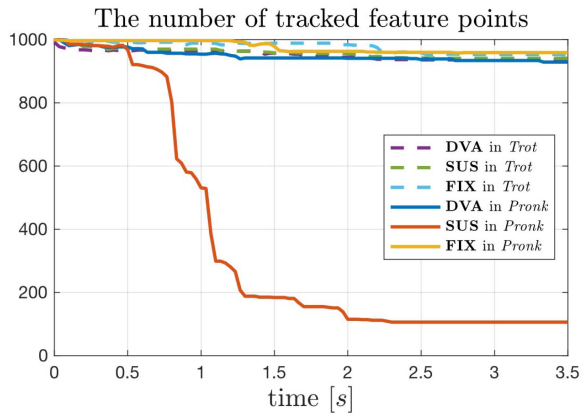


Fig. 13. Feature tracking performance with DVA, SUS, and FIX.

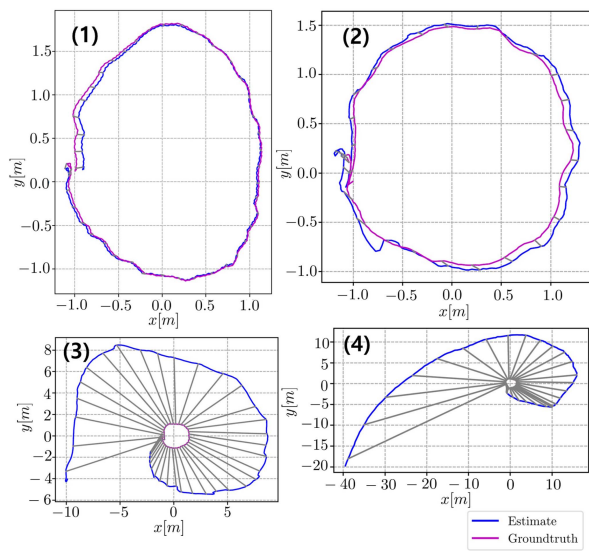


Fig. 14. The estimated trajectories of the robot and ground truth on the top view. These experiments are conducted in four cases: (1) DVA with trotting gait, (2) DVA with pronking gait, (3) FIX with trotting gait, and (4) FIX with pronking gait.

From Fig. 11, we can see the smooth head motions as compared to the body for the trotting gait with DVA and SUS. Also, we can see linear acceleration data showing a small impact and less noise attributed to the suspension mechanism compared to FIX for the trotting gait.

However, similar to the simulation results in Section V, for the case of pronking gait with SUS, the head exhibits unstable motion that rapidly moves with collision to its range limit as seen in Fig. 12. This resonance-like behavior is because SUS is not tunable, and results in the very noisy acceleration in Fig. 11 and poor feature tracking in Fig. 13, both significantly detrimental to the visual-inertial state estimation. On the other hand, for the case of FIX, the feature tracking in Fig. 13 is not perturbed much even during the pronking gait, yet, the impact and vibration directly transmitted to the IMU generate the noisy acceleration signal in Fig. 11. In contrast, the head with DVA maintains relatively small vertical motion, less noise in the

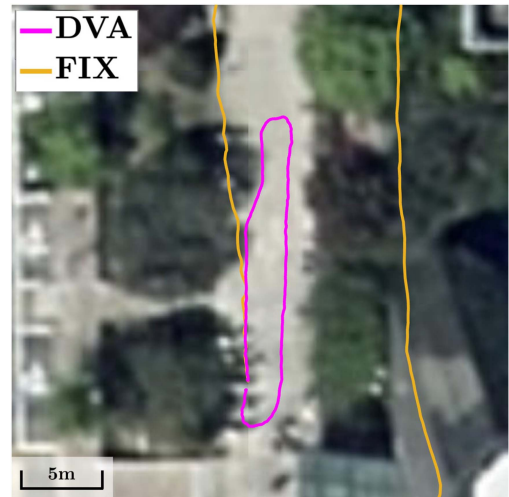


Fig. 15. The estimated trajectories of the outdoor experiment with DVA and FIX, aligned with aerial photography from NAVER map.

acceleration signal, and good feature tracking robustly across different gaits. A video of these experiments can be found at <https://youtu.be/xRF3SLvoh10>.

B. State Estimation Results

In Section VI-A, we verified that reliable sensor data can be measured with our proposed method. This section presents the state estimation results using these sensor data. We exploit VINS-MONO [27] for visual-inertial state estimation with a monocular camera and do not use loop closure to verify clearly how much drift is reduced.

1) *Indoor Experiment:* To verify the performance of the state estimation results, we use a MOCAP system as the ground truth and trajectory evaluation toolbox provided in [28]. In this experiment, we use the two necks (DVA and FIX) with two gaits (trot and pronk). We exclude SUS for this indoor and also the outdoor experiments (Section VI-B2), since its motion is so violent even to the verge of breaking the neck mechanism (see Figs. 11 and 12).

Fig. 14 shows the dramatically different estimated trajectories on the top view. In the case of using FIX, the result of trotting gait shows a fast drift, and the result of pronking gait even completely diverges, even though the visual perception pipeline appears adequate (e.g., good feature tracking in Fig. 13). This is in fact due to the effect of impact and vibration, directly landing on, and corrupting the acceleration signal of, the IMU, which is in particular crucial for the monocular visual-inertial state estimation, where the IMU is the sole information to correct scale ambiguity of monocular vision. Recall that mitigating this impact/vibration across different gaits cannot be attained with the non-tunable SUS and necessitates the tunable DVA. In contrast to the result of FIX, the case of utilizing the stable sensor data with DVA shows a slow drift. For the trotting gait, the root means square error for the position has 4.05 cm and for the pronking gait, 7.77 cm along the trajectories in Fig. 14. This clearly shows the efficacy of our proposed neck for dynamic

legged locomotion particularly only with a monocular camera and an IMU. A video of these experiments can be found at <https://youtu.be/O0T58hQ5gJE>.

2) *Outdoor Experiment*: In addition to indoor experiments, we compare **DVA** and **FIX** in an outdoor setting to verify our proposed method along a longer trajectory. In these experiments, the performance is confirmed by measuring the offset when returning to the starting position after taking a turn around the perimeter with a trotting gait.

The estimated trajectories are shown in Fig. 15, aligned with aerial photography to visually verify their accuracy. Similar to the result of indoor experiments, the case of using **FIX** shows a fast drift induced by the noisy IMU data. On the other hand, in the case of using **DVA**, its final drift without loop closure is $[-0.60, 0.05, -0.07]m$ in x, y, and z-axis, which occupies 1.50% of the total trajectory length. A video of these experiments can be found at <https://youtu.be/Dn7FNoyO624>.

VII. CONCLUSION

In this paper, we propose a tunable neck design for the robust state estimation of dynamic legged robots. Adjusting the neck to obtain stable sensor data with diverse locomotion enables a robust visual-inertial state estimation system. Mechanical design, modeling, and analysis of the proposed system are provided and the validity is verified through simulation and experiments.

This research has shown that mechanical neck design is indeed a promising solution to the problem of state estimation for dynamic legged robots. Our approach would also reduce the drift problem caused by slippages in the state estimation method using leg encoder information by providing more stable visual-inertial sensing. We will further refine the neck design with variable stiffness, lightweight, and small form factor. For example, with the help of the spring's fast response, we may also actively control the DVA to quickly stabilize the head even during transitions or non-periodic motions such as stepping down. We will also explore alternative directions to achieve variable stiffness of our neck instead of the current method of changing the inertial parameters of the dynamic equation.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of Biomimetic Robotics Lab at MIT and Naver Labs.

REFERENCES

- [1] Y. Lee, J. Yoon, H. Yang, C. Kim, and D. Lee, "Camera-GPS-IMU sensor fusion for autonomous flying," in *Proc. IEEE 8th Int. Conf. Ubiquitous Future Netw.*, 2016, pp. 85–88.
- [2] H. Yang, Y. Lee, S.-Y. Jeon, and D. Lee, "Multi-rotor drone tutorial: Systems, mechanics, control and state estimation," *Intell. Serv. Robot.*, vol. 10, no. 2, pp. 79–93, 2017.
- [3] Y. Lee, W. Do, H. Yoon, J. Heo, W. Lee, and D. Lee, "Visual-inertial hand motion tracking with robustness against occlusion, interference, and contact," *Sci. Robot.*, vol. 6, no. 58, 2021, Art. no. eabe1315.
- [4] R. Hartley, M. G. Jadidi, L. Gan, J.-K. Huang, J. W. Grizzle, and R. M. Eustice, "Hybrid contact preintegration for visual-inertial-contact state estimation using factor graphs," in *Proc. IEEE RSJ Int. Conf. Intell. Robots Syst.*, 2018, pp. 3783–3790.
- [5] D. Wisth, M. Camurri, and M. Fallon, "Robust legged robot state estimation using factor graph optimization," *IEEE Robot. Automat. Lett.*, vol. 4, no. 4, pp. 4507–4514, Oct. 2019.
- [6] Y. Kim, B. Yu, E. M. Lee, J.-H. Kim, H.-W. Park, and H. Myung, "Step: State estimator for legged robots using a preintegrated foot velocity factor," *IEEE Robot. Automat. Lett.*, vol. 7, no. 2, pp. 4456–4463, Apr. 2022.
- [7] M. Bloesch et al., "State estimation for legged robots: consistent fusion of leg kinematics and imu," *Robot.: Sci. Syst.*, vol. 17, pp. 17–24, 2013.
- [8] R. Hartley, M. Ghaffari, R. M. Eustice, and J. W. Grizzle, "Contact-aided invariant extended Kalman filtering for robot state estimation," *Int. J. Robot. Res.*, vol. 39, no. 4, pp. 402–430, 2020.
- [9] J.-H. Kim et al., "Legged robot state estimation with dynamic contact event information," *IEEE Robot. Automat. Lett.*, vol. 6, no. 4, pp. 6733–6740, Oct. 2021.
- [10] D. Wisth, M. Camurri, and M. Fallon, "VILENS: Visual, inertial, lidar, and leg odometry for all-terrain legged robots," 2021, *arXiv:2107.07243*.
- [11] Boston Dynamics, "Spot specification," 2016, Accessed: Jan. 30, 2023. [Online]. Available: <https://support.bostondynamics.com/s/article/Robot-specifications>
- [12] National Geographic, "Cheetah on edge," 2012, Accessed: Jan. 30, 2023. [Online]. Available: <https://www.nationalgeographic.com/magazine/article/cheetahs>
- [13] B. Katz, J. Di Carlo, and S. Kim, "Mini cheetah: A platform for pushing the limits of dynamic quadruped control," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2019, pp. 6295–6301.
- [14] O. Bruneau, F. B. Ouedzou, and J.-G. Fontaine, "Dynamic walk of a bipedal robot having flexible feet," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2001, pp. 512–517.
- [15] M. Hutter, C. Gehring, M. A. Höpflinger, M. Blösch, and R. Siegwart, "Toward combining speed, efficiency, versatility, and robustness in an autonomous quadruped," *IEEE Trans. Robot.*, vol. 30, no. 6, pp. 1427–1440, Dec. 2014.
- [16] N. G. Tsararakis, S. Morfeý, G. M. Cerda, L. Zhibin, and D. G. Caldwell, "Compliant humanoid coman: Optimal joint stiffness tuning for modal frequency control," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2013, pp. 673–678.
- [17] B. R. P. Singh and R. Featherstone, "Mechanical shock propagation reduction in robot legs," *IEEE Robot. Automat. Lett.*, vol. 5, no. 2, pp. 1183–1190, Apr. 2020.
- [18] H. Kumar, J. J. Payne, M. Travers, A. M. Johnson, and H. Choset, "Periodic SLAM: Using cyclic constraints to improve the performance of visual-inertial SLAM on legged robots," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2022, pp. 9477–9483.
- [19] S. Elias and V. Matsagar, "Research developments in vibration control of structures using passive tuned mass dampers," *Annu. Rev. Control.*, vol. 44, pp. 129–156, 2017.
- [20] H. Yao, T. Wang, B. Wen, and B. Qiu, "A tunable dynamic vibration absorber for unbalanced rotor system," *J. Mech. Sci. Technol.*, vol. 32, no. 4, pp. 1519–1528, 2018.
- [21] C. Laschi and B. Mazzolai, "Lessons from animals and plants: The symbiosis of morphological computation and soft robotics," *IEEE Robot. Automat. Mag.*, vol. 23, no. 3, pp. 107–114, Sep. 2016.
- [22] S. Nah, T. Hyun Kim, and K. Mu Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 3883–3891.
- [23] V. Klemm et al., "Ascento: A two-wheeled jumping robot," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2019, pp. 7515–7521.
- [24] G. A. Pratt and M. M. Williamson, "Series elastic actuators," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. Hum. Robot Interact. Cooperative Robots*, 1995, vol. 1, pp. 399–406.
- [25] D. Kim, J. Di Carlo, B. Katz, G. Bledt, and S. Kim, "Highly dynamic quadruped locomotion via whole-body impulse control and model predictive control," 2019, *arXiv:1909.06586*.
- [26] G. Bledt, P. M. Wensing, S. Ingersoll, and S. Kim, "Contact model fusion for event-based locomotion in unstructured terrains," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 4399–4406.
- [27] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1004–1020, Aug. 2018.
- [28] Z. Zhang and D. Scaramuzza, "A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2018, pp. 7244–7251.