Neck Design for Robust Visual-Inertial State Estimation during Aggressive Locomotion with Dynamic Vibration Absorber

Taekyun Kim and Dongjun Lee

***Abstract*— We propose a new neck design of legged robots to achieve robust visual-inertial state estimation in dynamic locomotion. Our neck design absorbs the impact from the legs and reduces the periodic vertical movement of the head using the dynamic vibration absorber (DVA), which can readily adapt to the frequency of oscillation of the body. Due to this smooth movement of the head, sensors in the head can measure data stably, which is impossible when sensors are directly attached to the body. We present the mechanical design of the neck as a combination of suspension mechanism with dynamic vibration absorber, which provide an adjustable notch filter for the vertical motion of the robot. Simulation and experimental results are performed to verify the effect of the proposed neck on the head of robots, manifesting that the states estimated from the visual-inertial sensors on the head can be precise even during aggressive motion by rendering both the inertial information and the feature tracking more stable and robust. The presented design can overcome the disadvantage that the localization of robots with legs can only operate well in gentle locomotion.**

***Paper Type* – Original Work**

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 a precise state estimation for safe autonomy. Leg odometry is widely used for that, utilizing the kinematic encoder data and inertial measurements. For this method, many approaches have been investigated based on filtering [1]–[3] and factor graph optimization [4], [5].

The drift issue of leg odometry is however inevitable because it relies only on the proprioceptive sensors: inertial measurement unit (IMU) and encoders. Therefore, many studies have recently focused on their fusions with visual information to reduce the drift [6], [7]. This approach is preferred since it only requires the addition of a camera, which is fairly economic both in price and its power budget. This state estimation method using visual-inertial measurements, however, is fragile during aggressive dynamic locomotions such as pronking and bounding gait, since: 1) the accelerometer signals become very noisy due to the body vibration and shock from the impact and 2) the camera may capture motion blur or detect only a small overlap between camera views, rendering the feature tracking challenging and even often diverging. For these reasons, some legged robots adopts many cameras and/or LiDAR sensor for their localization (e.g, five cameras for Spot [8], a camera with Velodyne LiDAR sensor for Anymal [9]), which result in high hardware and system integration costs.

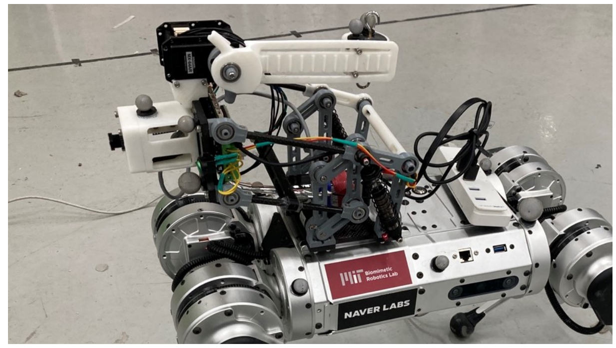

Fig. 1. The prototype of proposed neck design with a Mini Cheetah [10].

In contrast to these rather purely *sensor/software-oriented* approaches [6, 7, 9], here, we attempt to solve the perception problem of legged robots by recruiting a suitable mechanical hardware - a properly-designed vibration/impact-suppressing neck. This approach is in fact inspired by the nature, where many animals have not only eyes on the head, but also a neck connecting the head to the body arguably suitable for perception-stabilization. More specifically, we design our neck as a cascade of a suspension mechanism with a dynamic vibration absorber. The suspension mechanism provides suppression of vibration induced by the locomotion, absorbs the impact from the legs, and compensates the weight of the head consisting of a mono-camera and an IMU. We also adopt the concept of dynamic vibration absorber (DVA) for the neck design, which can provide a much faster (i.e., analog) response of notch filter as compared to active control with sampling rate, with its notching frequency also easily adjustable to the gait frequency on fly. This DVA mechanism

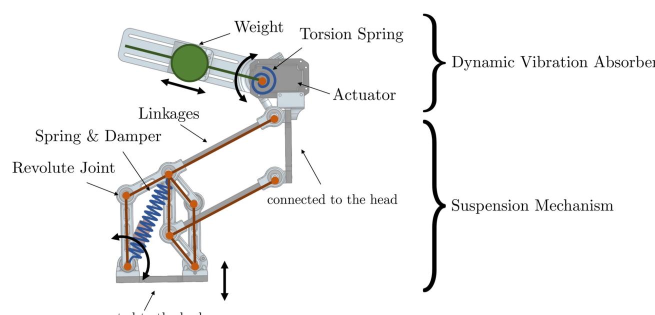

Fig. 2. The overall proposed neck design and its main components.

This work was supported by a Grant to Bio-Mimetic Robot Research Center funded by Defense Acquisition Program Administration, and by Agency for Defense Development of Korea (UD190018ID).

The authors are with the Department of Mechanical Engineering, IAMD and IER, Seoul National University, Seoul, Republic of Korea. {ktk4532, djlee}@snu.ac.kr, Corresponding author: Dongjun Lee

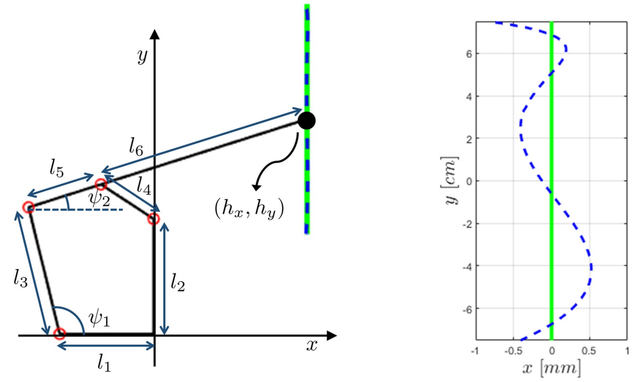

Fig. 3. (left) Kinematics model and design parameters of the proposed neck design. (right) The desired vertical trajectory and the trajectory of the head with the optimized design parameter. The dashed blue line indicates the trajectory of the head determined by the geometric parameters l_i. The green line denotes the desired vertical trajectory of the head.

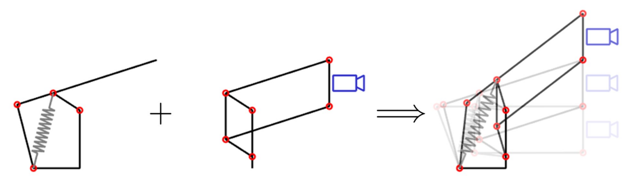

Fig. 4. The diagram of the suspension mechanism. The entire neck structure consists of a combination of a linkage generating vertical motion and two connected parallel four-bar linkage.

turns out to be crucial during certain aggressive locomotion, particularly those of low frequency pronking gaits. This is because, without the DVA, the head can resonate with the locomotion and reach its range limit, thereby, compromising the accelerometer signal of the IMU, which is critical for the proper working of visual-inertial state estimation. In contrast, with the DVA, such excessive head motion can be channeled from the head motion to the DVA motion. This perception-stabilizing efficacy of our DVA-based neck design is also verified with simulations and experiments using MIT mini cheetah.

II. Mechanical Design

As seen in Fig. 2, the proposed neck design is mainly composed of the two parts: 1) the suspension mechanism that compensates the weight of the head, absorbs the impact from the legs, and reduces the vibration of the head; and 2) the adjustable DVA that absorbs the periodic excessive motion of the head.

A. Linkage Design for Suspension Mechanism

The main parts of the suspension consist of a linkage mechanism that connects the body and head. This linkage is designed for two goals: providing the vertical head motion as much as possible, and rendering the direction of the head the same as the body. For the vertical movement of the head, we design a linkage which is a modification of the leg design of Ascento [11] in a way that places the trajectory of the linkage slightly forward so as not to block the view of the camera, as seen in Fig. 3. This linkage mechanism cannot make the

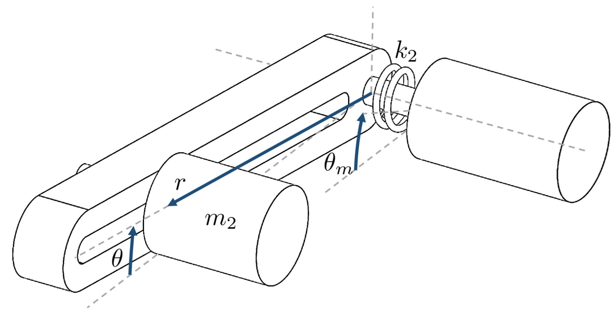

Fig. 5. The diagram of adjustable dynamic vibration absorber (DVA). By changing the weight position along with the slider, it can adjust the natural frequency.

perfectly vertical movement of the head 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.

Moreover, we supplement this vertically moving linkage with the two connected parallel four-bar linkage as Fig. 4. 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.

B. Adjustable Dynamic Vibration Absorber

Although the above-mentioned suspension mechanism can reduce vibration to some extent, we developed an adjustable DVA to handle large motions of low frequencies. This system is attached to the suspension mechanism, seen in Fig. 2, absorbing the dynamic oscillation of the head.

The proposed DVA is composed of an actuator, a torsion spring, a slider, and a weight. A torsion spring connects the actuator to the slider, similar to series elastic actuators, as shown in Fig. 5. By changing the position of weight along with the slider, the notch frequency of the system is determined and then the actuator rotates the rod to maintain it horizontally in steady state.

III. Simulation

To verify our proposed method, we conducted simulations with the original dynamics of the overall system. The simulations were performed with two sinusoid input models similar to the data of real robot motion, seen in Table I.

TABLE I

Input Models as Vertical Motion in Simulation

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

Fig. 6 shows the simulation results about the head motion with respect to the body motion. In the case of not using DVA, the head position without DVA also show the stable head motion for the trot input model. This is because the trot input has high frequency and small amplitude, so only

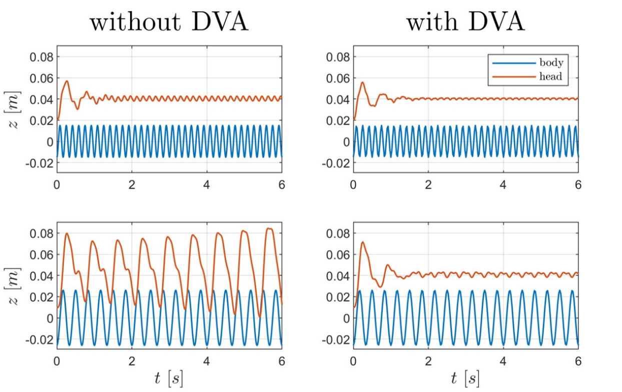

Fig. 6. The z position of body and head in simulation (left) without DVA and (right) with DVA. The simulations were conducted in two input models: (top) trot gait and (bottom) pronking gait.

a suspension mechanism can alleviate the vibration of the head. On the other side, for the case of pronking gait, the head motions become unstable, as it resonates with the body movement due to the nonlinear terms neglected when simplifying. In contrast, it was confirmed that DVA induces more stable head movement in the neck because it absorbs large oscillation even with challenging environments.

IV. Experiments

In this section, experimental results with the proposed prototype of the neck design are presented. Each experiment was conducted with a Mini Cheetah [10] which has strength in dynamic locomotion. For demonstrating the effect of the proposed neck mechanism on the state estimation of dynamic legged robots, the contents of the experiment are as follows.

- Vertical movement of the head for the body movement (IV-A).
- The linear acceleration data in the head (IV-B).
- Localization results with the proposed neck (IV-C).

Every experiment is performed with trotting gait and pronking gait that represent the gentle and dynamic locomotion, respectively.

A. Stabilization of the vertical movement

One of the goals to use the neck mechanism is stabilizing the periodic vertical movement of legged robots. To validate the effect of the suspension mechanism and DVA on the vertical movement, the OptiTrack MOCAP (motion capture) system measured the motions of the head and body with 100Hz. These experiments were conducted with and without DVA.

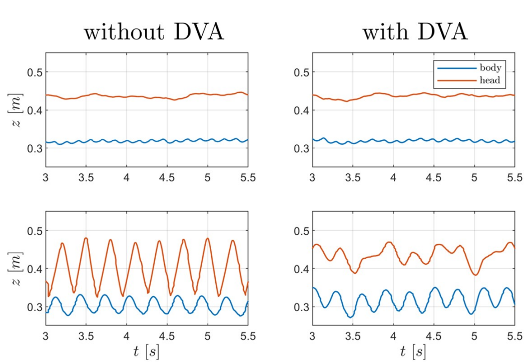

Fig. 7. The z position data from MOCAP system without DVA (left) and with DVA (right). These experiments were conducted in two gaits: trot (top) and pronk (bottom).

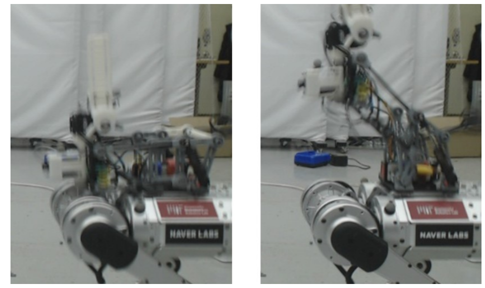

Fig. 8. Two snapshots showing the collision of the neck, in the case of pronking gait without DVA.

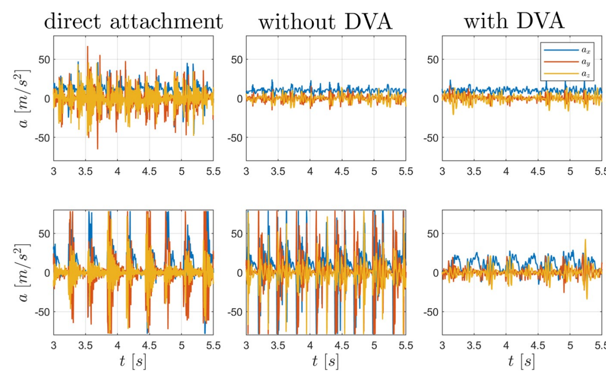

Fig. 9. Linear acceleration data measured by an IMU in the heads (left) directly attached to the body, (middle) neck mechanism without DVA, and (right) with DVA. These experiments were conducted in two gaits: (top) trot and (bottom) pronk.

Fig. 7 shows the data of the experiments, from which we can see the smooth head motions as compared to the body in trotting gait, whether using DVA or not. It allows the camera to take more stable images because the head vertically moves slower than the body. However, if not use the DVA in a pronking gait, the head shows unstable motion. that shows sharp moving with contact to its range limit as seen in Fig. 8. Although there are still some fluctuations, the head with DVA has a much softer movement than the case without DVA. Through this stable motion of the head, the camera can capture smooth images that enhance the performance of feature tracking used for state estimation.

B. Stabilization of the IMU data

Reliable linear acceleration data is also critical for robust visual-inertial state estimation. The accelerometer data were measured in three cases: 1) directly attached to the body, 2) on the head without DVA, and 3) on the head with the overall neck mechanism. From Fig. 9, we can see stable linear acceleration data attributed to the suspension mechanism in trotting gait. Similar to results in Sec. IV-A, without the DVA in pronk motion, it collides with the range limit and thereby shows very noisy acceleration data. Conversely, acceleration

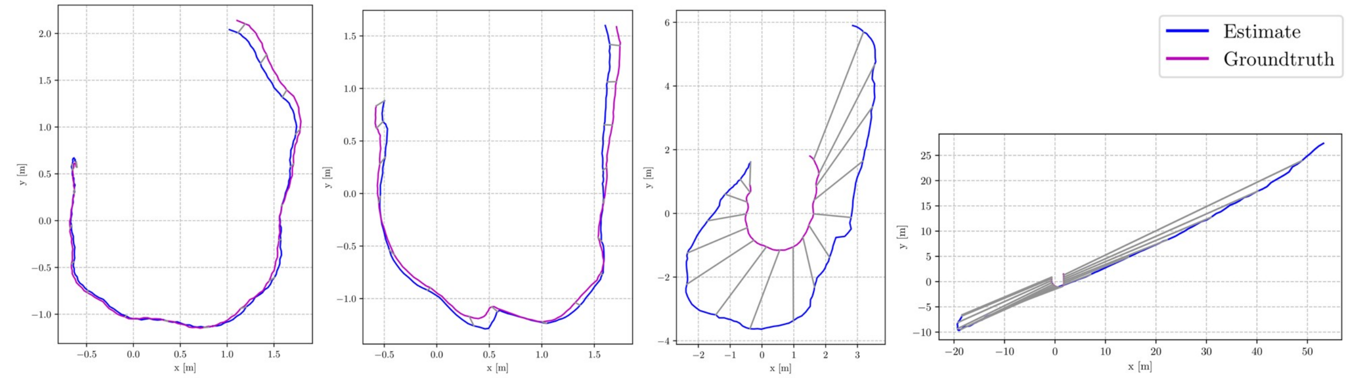

Fig. 10. The estimated trajectories of the robot and ground truth on the top view. This experiment was conducted in four cases: (first) trot gait with the neck, (second) pronking gait with the neck, (third) trot gait without the neck, and (fourth) pronking gait without the neck.

data using DVA shows less noise and small fluctuations due to the oscillation of the head.

C. State Estimation Results

In Sec. IV-A, B, we verified that the stable data can be measured with our proposed method. This section presents the state estimation results using these stable sensor data.

1) Experiment Setup: We exploited VINS-MONO [12] for visual-inertial based state estimation with a mono-camera. We use a MOCAP system as ground truth to verify the performance of this VIO algorithm with the proposed neck and trajectory evaluation toolbox, provided in [13].

2) Estimated Trajectory: Fig. 10 shows the dramatically different estimated trajectory on the top view. In the case of not using the neck, the trotting gait showed a fast drift, and the pronking gait even completely diverged. It originated from uncertain landmark observation due to shaking camera movement and very noisy acceleration data due to impact and vibration, seen in Sec. IV-B. On the other hand, the case of utilizing the stable sensor data through the neck shows a slow drift. For trotting gait, the root means square error for the position has $6.48cm$ and for pronking gait, $7.28cm$ along the trajectories in Fig. 10.

V. Conclusions

In this paper, we proposed a new neck design for the robust state estimation of legged robots. By obtaining stable inertial and visual information in dynamic motion, it enables a robust visual-inertial state estimation system in dynamic locomotion. Mechanical design, modeling, and analysis of the proposed system were provided and the validity was verified through simulation and experiments.

This research has shown that mechanical neck design is indeed a feasible solution to the problem of state estimation for dynamic legged robots. We will later develop the neck design in a way that increases the stability against motion, vibration, and impact in varying locomotion. Also, we will fuse this stable visual-inertial information with the kinematic information of the legs and neck for improving the localization of dynamic legged robots.

Acknowledgment

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

References

[1] M. Bloesch, M. Hutter, M. Hoepflinger, S. Leutenegger, C. Gehring, C. D. Remy *et al.*, "State estimation for legged robotsconsistent fusion of leg kinematics and imu," *Robotics: Science and Systems*, vol. 17, pp. 17–24, 2013.

[2] M. Bloesch, M. Burri, H. Sommer, R. Siegwart, and M. Hutter, "The two-state implicit filter recursive estimation for mobile robots," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 573–580, 2017.

[3] R. Hartley, M. Ghaffari, R. M. Eustice, and J. W. Grizzle, "Contact-aided invariant extended kalman filtering for robot state estimation," *The International Journal of Robotics Research*, vol. 39, no. 4, pp. 402–430, 2020.

[4] R. Hartley, J. Mangelson, L. Gan, M. G. Jadidi, J. M. Walls, R. M. Eustice, and J. W. Grizzle, "Legged robot state-estimation through combined forward kinematic and preintegrated contact factors," in *IEEE International Conference on Robotics and Automation*. IEEE, 2018, pp. 4422–4429.

[5] J.-H. Kim, S. Hong, G. Ji, S. Jeon, J. Hwangbo, J.-H. Oh, and H.-W. Park, "Legged robot state estimation with dynamic contact event information," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 6733–6740, 2021.

[6] 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 *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems)*. IEEE, 2018, pp. 3783–3790.

[7] D. Wisth, M. Camurri, and M. Fallon, "Robust legged robot state estimation using factor graph optimization," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4507–4514, 2019.

[8] Boston Dynamics, "Spot specification," https://support.bostondynamics.com/s/article/Robot-specifications, 2016, accessed on 2022-03-07.

[9] D. Wisth, M. Camurri, and M. Fallon, "Vilens: Visual, inertial, lidar, and leg odometry for all-terrain legged robots," *arXiv preprint arXiv:2107.07243*, 2021.

[10] B. Katz, J. D. Carlo, and S. Kim, "Mini cheetah: A platform for pushing the limits of dynamic quadruped control," in *IEEE International Conference on Robotics and Automation*. IEEE, 2019, pp. 6295–6301.

[11] V. Klemm *et al.*, "Ascento: A two-wheeled jumping robot," in *2019 International Conference on Robotics and Automation*. IEEE, 2019, pp. 7515–7521.

[12] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.

[13] Z. Zhang and D. Scaramuzza, "A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2018, pp. 7244–7251.