Textile & Leather Review

ISSN 2623-6281 | www.tlr-journal.com | 10.31881/TLR

Active Vibration Suppression and Trajectory Tracking Control of a Precision Motion Stage via Deep Learning Based Multisensor Fusion

Zeyu Gu

How to cite: Gu Z. Active Vibration Suppression and Trajectory Tracking Control of a Precision Motion Stage via Deep Learning Based Multisensor Fusion. Textile & Leather Review. 2026; 9:2892-2921. https://doi.org/10.31881/TLR.2026.2892

How to link https://doi.org/10.31881/TLR.2026.2892

Published:25 April 2026

Active Vibration Suppression and Trajectory Tracking Control of a Precision Motion Stage via Deep Learning Based Multi-sensor Fusion

Zeyu Gu

School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China
g15525081680@163.com

Article
https://doi.org/10.31881/TLR.2026.2892
Published 25 April 2026

ABSTRACT

To address the challenge of simultaneously achieving high-precision trajectory tracking and active vibration suppression in precision motion stages under complex dynamic operating conditions, this paper proposes a deep learning-based multisensor fusion control method. By exploiting the complementary measurement characteristics of position and acceleration sensors, a multisensor fusion model is constructed for dynamic state estimation, allowing control-relevant features related to motion states, lumped disturbances, and vibration evolution to be extracted from multisource temporal data. Based on a baseline feedback controller, a fusion-estimation-driven compensation strategy is then designed to incorporate disturbance compensation and vibration suppression into the closed-loop framework, thereby improving tracking accuracy and dynamic smoothness. Experiments, including standard trajectory tracking, disturbance rejection, and high-speed dynamic tests, are conducted on a single-axis precision motion stage and used to compare the proposed method with a baseline feedback controller and a conventional fusion/observer-enhanced method. Results show that the proposed method achieves lower tracking errors, smaller vibration peaks, faster residual vibration decay, and better disturbance tolerance than the comparison methods. These findings indicate that the proposed fusion-control framework is effective for improving the overall performance of precision motion systems.

KEYWORDS
precision motion stage, multisensor fusion, active vibration suppression, trajectory tracking

---

INTRODUCTION

High-performance precision motion stages are widely used in advanced manufacturing applications such as semiconductor fabrication, precision machining, optical alignment, and micro/nano-manipulation[1, 2]. The motion accuracy and dynamic response of these systems directly determine machining quality, positioning precision, and operational efficiency[3, 4]. As industrial systems continue to evolve toward higher speed, higher acceleration, and higher precision, precision motion stages are increasingly subject to structural flexibility, external disturbances, unmodeled dynamics, and measurement noise during rapid start-stop maneuvers and complex trajectory tracking tasks[5]. These factors often lead to residual vibration, enlarged tracking errors, and even overall performance degradation[6, 7]. Under high-bandwidth control requirements, achieving effective vibration suppression while maintaining fast response and high tracking accuracy has therefore become a critical challenge in the field of precision motion control[8, 9].

To address these issues, extensive efforts have been devoted to trajectory tracking and vibration suppression for precision motion stages, leading to the development of a variety of control strategies, including proportional-integral-derivative control, disturbance observers, repetitive control, sliding mode control, model predictive control, and input shaping. Although these methods have improved system performance to some extent, most of them rely heavily on relatively accurate dynamic models and may suffer from degraded performance in the presence of strong nonlinearities, time-varying disturbances, and parameter uncertainties[10, 11]. Meanwhile, multisensor fusion has attracted increasing attention because it can integrate heterogeneous information from position, velocity, and acceleration measurements to enhance state awareness and noise robustness, and has shown promising potential in precision positioning and vibration control applications[12]. However, conventional fusion approaches are still largely built upon linear assumptions or model-based formulations, making it difficult to fully exploit the latent nonlinear and temporal correlations among multisource measurement signals. In recent years, deep learning has demonstrated remarkable capabilities in dynamic system modeling, feature extraction, and sequence prediction, offering a new avenue for multisensor fusion and intelligent control in precision motion systems[13, 14]. Nevertheless, integrated studies that simultaneously address active vibration suppression and high-precision trajectory tracking for precision motion stages remain relatively limited.

Motivated by these considerations, this paper proposes a deep learning-based multisensor fusion control approach for active vibration suppression and high-precision trajectory tracking of a precision motion stage under complex dynamic operating conditions. First, a deep learning-driven multisensor fusion model is developed by leveraging multisource measurements from the motion stage to enhance the representation of dynamic states and disturbance-related information. Then, a compensation control strategy is designed based on the fused estimation results and integrated with a baseline feedback controller to simultaneously improve vibration suppression and trajectory tracking performance. Finally, comparative experiments are conducted to validate the effectiveness of the proposed method in terms of tracking accuracy, vibration attenuation, and robustness against conventional approaches. The present study provides a new technical framework for intelligent control and multisensor collaborative perception in precision motion systems.

SYSTEM DESCRIPTION AND PROBLEM FORMULATION

Precision Motion Stage and Sensor Configuration

The system considered in this study is a typical single-axis precision motion stage consisting of an actuation unit, a moving platform, a guiding and supporting structure, a real-time controller, and multiple sensors. Such systems commonly operate under high-speed, high-acceleration, and high-precision conditions, and are widely used in precision positioning, micro/nano-manufacturing, and advanced industrial equipment. In practical operation, the motion stage is required not only to track the prescribed reference trajectory rapidly and accurately, but also to suppress vibrations induced by structural flexibility, external excitation, and actuator dynamics so as to maintain desirable dynamic performance and positioning accuracy[15].

To achieve comprehensive perception of both motion and vibration characteristics, a multisensor measurement scheme is adopted in this work. Specifically, the position sensor is used to acquire the displacement of the stage and typically provides high static measurement accuracy as well as reliable low-frequency response. The accelerometer is employed to capture high-frequency vibration and dynamic disturbances, thereby reflecting rapidly varying local dynamic features of the platform. When available, auxiliary measurements such as velocity information or drive current can also be incorporated to further enhance the representation of the system dynamics. Since different sensors exhibit distinct characteristics in terms of measurement bandwidth, noise level, accuracy, and dynamic sensitivity, a single sensor is often insufficient to simultaneously satisfy the requirements of high-precision positioning and high-frequency vibration perception. Therefore, constructing a state perception mechanism capable of integrating heterogeneous multisource information is a critical prerequisite for high-performance motion control.

From the perspective of control architecture, the precision motion stage can be described as a closed-loop system composed of a reference input, a controller, an actuator, the motion plant, and a measurement module. The reference trajectory is fed into the controller to generate the drive signal, which is then applied to the stage through the actuator to realize the desired motion. Meanwhile, the multisensor system continuously collects displacement, vibration, and other dynamic response information from the platform and feeds them back to the controller or the fusion module for state estimation and compensation design. To facilitate the development of the proposed approach, a dynamic model that explicitly accounts for vibration modes, external disturbances, and measurement noise is established in the following subsection, while multisensor fusion is incorporated into the overall framework as an essential component for enhanced state perception.

Dynamic Modeling and Disturbance Analysis

Considering the structural flexibility and disturbance-coupling effects encountered during practical operation, the precision motion stage is modeled here as a dynamic system composed of rigid-body motion superimposed with flexible vibration. Let $x(t)$ denote the actual displacement of the stage, $r(t)$ the reference trajectory, and $u(t)$ the control input. The system can be generally expressed as:

$$M\ddot{x}(t) + C\dot{x}(t) + Kx(t) + f_{nl}(t) + d(t) = Bu(t) \tag{1}$$

where M , C , and K denote the equivalent mass, damping, and stiffness parameters of the system, respectively; $f_{nl}(t)$ represents nonlinear effects such as friction, hysteresis, dead-zone behavior, and unmodeled flexible dynamics; $d(t)$ denotes the lumped disturbance induced by external excitation and parameter uncertainty; and B is the input distribution matrix. For high-performance precision motion stages, the above nonlinearities and disturbances can significantly deteriorate trajectory tracking accuracy and easily excite structural resonance and residual vibration, especially under rapid start-stop maneuvers, acceleration/deceleration switching, and high-frequency reference inputs.

For the convenience of controller development and state analysis, the system can be further represented in state-space form. Define the state vector as:

$$\mathbf{x}(t) = [x(t), \dot{x}(t), z_1(t), z_2(t), ..., z_n(t)]^T \tag{2}$$

where $x(t)$ and $\dot{x}(t)$ denote the stage displacement and velocity, respectively, and $z_i(t)$ represents the extended states associated with dominant flexible modes or vibration dynamics. The system can then be abstractly written as:

$$\dot{\mathbf{x}}(t) = A\mathbf{x}(t) + Bu(t) + Ed(t) + \Phi(\mathbf{x}, t) \tag{3}$$

$$\mathbf{y}(t) = H\mathbf{x}(t) + \mathbf{v}(t) \tag{4}$$

where $\mathbf{y}(t)$ is the sensor output vector, H is the measurement matrix, $\mathbf{v}(t)$ denotes measurement noise, and $\Phi(\mathbf{x}, t)$ captures unmodeled nonlinear and time-varying dynamics. Because position sensors and accelerometers have fundamentally different observation characteristics, the resulting measurements typically contain useful information in different frequency ranges, while also being subject to their own noise and error propagation mechanisms. For instance, position measurements are more suitable for reflecting low-frequency displacement variations, whereas acceleration measurements are more effective in characterizing high-frequency vibration components, although direct integration of acceleration signals may lead to drift and accumulated noise. Accordingly, the extraction of control-relevant dynamic information from multisource measurements is a key factor in improving overall control performance.

Furthermore, the disturbances affecting the precision motion stage can be broadly categorized into three types. The first type consists of internal dynamic disturbances caused by structural flexibility, transmission imperfections, and nonideal actuator characteristics. The second type includes external disturbances arising from environmental vibration, load variation, and external impact. The third type is composed of measurement disturbances introduced by sensor noise, bias drift, and sampling errors. These disturbances are typically nonlinear, time-varying, and broadband in nature, which makes it difficult for conventional model-dependent control methods to achieve satisfactory comprehensive performance under complex operating conditions. Therefore, it is necessary to introduce a fusion mechanism capable of learning complex dynamic mappings from multisensor measurements so as to obtain more informative representations of system states and disturbances, thereby supporting active vibration suppression and trajectory tracking control in the subsequent design.

Control Objectives

In view of the above system characteristics and disturbance conditions, the control objectives of this study can be summarized in three aspects. First, for a given reference trajectory $r(t)$, the system is expected to achieve high-precision trajectory tracking such that the actual output $x(t)$ follows the desired motion as rapidly and accurately as possible. Accordingly, the tracking error is defined as:

$$e(t) = r(t) - x(t) \tag{5}$$

which should be maintained within a sufficiently small range throughout operation. Ideally, the controller should reduce overshoot, steady-state error, and dynamic tracking error while preserving fast transient response, so as to satisfy the requirements of precision positioning and complex motion execution.

Second, the system should possess effective active vibration suppression capability. In particular, under rapid acceleration/deceleration, abrupt trajectory variation, and external disturbance excitation, both the vibration amplitude and the residual vibration duration should be minimized as much as possible. Unlike pure position tracking, vibration suppression is more concerned with high-frequency dynamic response and flexible mode excitation. This requires the controller to promptly perceive and suppress vibration propagation and amplification caused by multiple dynamic factors. In other words, this work focuses not only on low-frequency positioning accuracy but also on high-frequency dynamic smoothness.

Finally, the control system should maintain satisfactory robustness and practical implementability in the presence of complex disturbances and measurement uncertainty. More specifically, the proposed method is expected to preserve favorable tracking and vibration suppression performance even when the system model is imperfect, the disturbances vary over time, and the sensor measurements are contaminated by noise. In addition, since precision motion stages are typically operated in real-time control environments, the developed fusion and control strategy should also exhibit adequate online computational efficiency and engineering feasibility. Based on these objectives, a deep learning-driven multisensor fusion model will be developed in the next section to enhance the perception of system states and disturbances, upon which an integrated control strategy for active vibration suppression and trajectory tracking will be established.

DEEP LEARNING-BASED MULTISENSOR FUSION FOR DYNAMIC STATE ESTIMATION

Characteristics of Multisensor Measurements

In precision motion control systems, different types of sensors exhibit substantially different capabilities in capturing motion states and vibration characteristics. These differences are reflected not only in the complementarity of the information they provide, but also in the heterogeneity of their error characteristics. Position sensors generally offer high-resolution displacement measurements with strong accuracy and stability in the low-frequency range, making them suitable for observing steady-state position, low-frequency motion trends, and trajectory tracking errors. However, due to their limited dynamic bandwidth, position sensors are often insufficient for fully characterizing high-frequency vibration components, especially during rapid start-stop operations, aggressive acceleration/deceleration transitions, and situations where flexible modes are excited.

In contrast, accelerometers are particularly effective in sensing high-frequency dynamics. They can capture platform vibration, impact response, and rapidly varying local dynamic behavior, and are therefore well suited for describing vibration propagation and flexible mode response in precision motion stages. Nevertheless, acceleration measurements also suffer from inherent limitations, such as increased sensitivity to measurement noise, weaker low-frequency stability, and drift accumulation when reconstructing velocity or displacement through direct integration. In some systems, auxiliary measurements such as velocity, drive current, actuation force, or strain may also be available. Although such signals do not directly represent displacement output, they can reveal actuator dynamics, load variation, and structural response from different perspectives, thereby providing additional support for dynamic state perception.

These observations indicate that the dynamic behavior of a precision motion stage cannot be fully characterized by any single sensor alone. Instead, a more comprehensive and robust state representation requires the integration of multisource measurements. Conventional fusion approaches typically combine heterogeneous signals through weighted averaging, complementary filtering, or model-based state estimation. However, these methods often rely on predefined model structures or linear assumptions, which limits their adaptability to nonlinear coupling, time-varying disturbances, and broadband dynamic responses. To address this issue, a deep learning-driven multisensor fusion mechanism is introduced in this study. By jointly modeling multisource temporal measurements, the proposed approach is able to extract latent dynamic correlations and temporal dependencies across sensors, thereby yielding dynamic state estimates that are more informative and more relevant to control.

Fusion Network Architecture

To fully exploit the temporal dependency and cross-modal complementarity embedded in multisensor measurements, a deep learning-based fusion network is developed for dynamic state estimation of the precision motion stage. The network takes multisource sensor sequences collected over consecutive sampling instants as input and outputs control-oriented dynamic state estimates through temporal feature extraction and information fusion. Since the precision motion stage exhibits pronounced dynamic memory, and the evolution of vibration propagation, disturbance accumulation, and tracking error is closely related to historical states, a gated recurrent unit (GRU)-based temporal modeling structure is adopted in this work to achieve a balance between modeling capability and real-time implementability.

Let the measurement vector at time step k , composed of position, acceleration, and auxiliary sensor signals, be denoted by:

$$\mathbf{s}_k = [p_k, a_k, \xi_k]^T \tag{6}$$

where p_k is the position measurement, a_k is the acceleration measurement, and ξ_k represents optional auxiliary measurements. Over a temporal window of length L , the network input is formulated as:

$$\mathbf{S}_k = [\mathbf{s}_{k-L+1}, \mathbf{s}_{k-L+2}, \ldots, \mathbf{s}_k] \tag{7}$$

This input sequence is first normalized and temporally organized so as to alleviate the influence of inconsistent physical dimensions and numerical ranges on the training process. The multisource sequence is then fed into the GRU encoder, which extracts latent features related to motion tendency, vibration evolution, and disturbance propagation. By recursively updating the hidden state at each time step, the GRU can selectively preserve useful historical information and suppress irrelevant perturbations through its gating mechanism, making it particularly suitable for modeling the time-varying dynamics of precision motion stages.

Following the GRU encoding layer, a feature mapping and fusion output layer is introduced to transform the temporal hidden states into a unified dynamic representation and to generate the final estimation output. In accordance with the control objectives of this work, the output of the fusion network is defined as an enhanced dynamic state estimation vector:

$$\hat{\mathbf{z}}_k = [\hat{x}_k, \hat{\dot{x}}_k, \hat{d}_k, \hat{v}_k]^T \tag{8}$$

where $\hat{x}_k$ and $\hat{\dot{x}}_k$ denote the fused estimates of position and velocity, respectively; $\hat{d}_k$ represents the estimate of the lumped disturbance or equivalent dynamic uncertainty; and $\hat{v}_k$ denotes the vibration-related latent state or vibration intensity descriptor. Compared with conventional control schemes that directly rely on a single measured output, the proposed fusion architecture jointly extracts low-frequency displacement information, high-frequency vibration features, and auxiliary dynamic information from the temporal input window, thereby providing the controller with a more complete and more robust state basis.

In the present implementation, the fusion network adopts a lightweight stacked-GRU architecture with two GRU layers. Each GRU layer has 64 hidden units, which was selected as a compromise between temporal modeling capability and real-time computational cost. The output hidden feature at the final time step is passed through a fully connected layer to generate the enhanced dynamic state estimation vector. A rectified linear unit (ReLU) activation function is used in the fully connected feature-mapping stage, while the GRU itself uses its standard gating nonlinearities, namely the sigmoid and hyperbolic tangent functions. The input sequence length is set to 20 sampling steps. Therefore, the network input dimension is determined by the number of sensor channels over the temporal window, and the output dimension is four, corresponding to the estimates of position, velocity, lumped disturbance, and vibration-related state.

From an engineering implementation perspective, the proposed fusion network is deliberately designed in a lightweight manner, avoiding excessively deep or overly complicated model stacking so as to reduce online computational burden and satisfy real-time control requirements. Moreover, the network is not intended to replace the conventional feedback controller. Instead, it serves as an enhanced state perception module embedded in the control loop to support active vibration suppression and trajectory tracking compensation. Such an architecture is not only compatible with existing control frameworks, but also more favorable for maintaining closed-loop stability and interpretability.

Network Training and Output Definition

To ensure that the fusion network output can faithfully represent the key dynamic states of the precision motion stage and directly support the subsequent controller design, a supervised learning strategy is adopted to establish the mapping between the multisensor inputs and the desired output targets. Specifically, multi-sensor synchronized data are collected under various operating conditions to construct a training dataset containing reference trajectories, position responses, acceleration responses, and auxiliary measurements. Based on system observations, offline identification results, or high-precision reference signals, target labels are then generated for network training. The objective of the network is not merely to fit a particular sensor signal, but rather to learn the coupling relationships among multisource temporal measurements so as to obtain fused estimates that are more representative of platform dynamics, disturbance variations, and vibration behavior.

Let the training target corresponding to the k -th sample be defined as:

$$\mathbf{z}_k = [x_k, \dot{x}_k, d_k, v_k]^T \tag{9}$$

where x_k , $\dot{x}_k$, d_k , and v_k denote the true or reference values of position, velocity, lumped disturbance, and vibration-related state, respectively. The network parameters are optimized by minimizing the following loss function:

$$\mathcal{L}=\frac{1}{N}\sum_{k=1}^{N}\left(\parallel \hat{\mathbf{z}}_k-\mathbf{z}_k \parallel_2^2+\lambda_1 \parallel \hat{d}_k-d_k \parallel_2^2+\lambda_2 \parallel \hat{v}_k-v_k \parallel_2^2\right) \tag{10}$$

where N is the number of training samples, and λ_1 and λ_2 are weighting coefficients used to regulate the relative importance of disturbance estimation and vibration-state estimation in the overall optimization objective. Built upon the global state estimation error, this loss function further strengthens the learning of disturbance- and vibration-related information, allowing the network output to better align with the integrated task of active vibration suppression and trajectory tracking.

For reproducibility, the network training procedure is specified as follows. The model is trained offline using the Adam optimizer with an initial learning rate of 1×10^{-3} and a batch size of 128. The training process runs for 100 epochs, and the model corresponding to the lowest validation loss is selected for deployment. The dataset is divided into training, validation, and test subsets in a ratio of 70%:15%:15%. To reduce temporal leakage, the split is performed on the basis of experimental runs rather than by randomly shuffling individual time samples, so that data from the same run do not simultaneously appear in different subsets. In addition, all input channels are normalized using the mean and standard deviation computed from the training set, and the same normalization statistics are applied to the validation and test sets.

In terms of output definition, the proposed fusion network is regarded as an enhanced dynamic state estimator rather than a conventional signal filter. Its core role is to provide the controller with dynamically informative state inputs, especially those disturbance- and vibration-related quantities that are difficult to acquire reliably from any single sensor. Based on this formulation, the subsequent controller can utilize $\hat{d}_k$ to construct compensation terms that mitigate the influence of lumped disturbances, while $\hat{v}_k$ can be used to characterize vibration evolution and thereby strengthen the coordination between trajectory tracking and vibration suppression.

It should be emphasized that the proposed fusion network does not aim to reconstruct all internal physical states of the system in a strict sense. Instead, it focuses on control-oriented state enhancement, namely, learning those dynamic features that are most critical to improving closed-loop performance. Through this task-oriented output definition, the fusion model can remain structurally compact and computationally efficient while serving the subsequent controller design more directly. This also lays the foundation for the integrated control strategy for active vibration suppression and trajectory tracking to be presented in the next section.

INTEGRATED CONTROL DESIGN FOR ACTIVE VIBRATION SUPPRESSION AND TRAJECTORY TRACKING

Baseline Feedback Controller

To ensure the basic closed-loop stability and nominal tracking capability of the precision motion stage, a baseline feedback controller is first established as the core stabilizing component of the overall control framework. Considering that practical precision motion systems generally require a control structure that is transparent, tunable, and suitable for real-time implementation, a composite feedback controller based on position error and velocity feedback is adopted in this work as the baseline controller. Its control law can be expressed as:

$$u_b(k) = K_p e(k) + K_d \dot{e}(k) + K_i \sum_{j=0}^{k} e(j) \tag{11}$$

where $u_b(k)$ denotes the baseline control input, $e(k) = r(k) - x(k)$ is the trajectory tracking error, and K_p , K_d , and K_i are the proportional, derivative, and integral gains, respectively. The main role of this controller is to provide fundamental closed-loop regulation capability, allowing the system to achieve nominal trajectory tracking under general operating conditions while suppressing low-frequency errors and steady-state deviations.

Although conventional feedback controllers are highly practical in precision motion control, their performance is often significantly affected by model uncertainty, flexible vibration, external disturbances, and measurement noise. In particular, under high-speed and high-acceleration operating conditions, trajectory error and vibration response are usually strongly coupled. On the one hand, structural vibration directly deteriorates transient tracking performance; on the other hand, increasing the feedback gain to achieve faster response may further excite high-frequency modes, thereby creating an inherent conflict between rapid tracking and effective vibration suppression. In view of this issue, the present study does not simply rely on increasing the controller bandwidth to improve performance. Instead, while preserving the stabilizing role of the baseline feedback controller, a fusion-estimation-driven enhancement mechanism is introduced to compensate for complex disturbances and vibration dynamics in a targeted manner.

Structurally, The baseline feedback controller serves two important functions in the proposed framework. First, it acts as the main control loop to guarantee the basic stability and tracking convergence of the system. Second, it provides a stable and controllable embedding platform for the subsequent deep learning-driven compensation module. In other words, the deep learning fusion module is not intended to replace the conventional controller, but rather to work cooperatively with it in the form of an additional compensation term so as to enhance the system's adaptability to complex dynamic conditions. To further substantiate our claim that the proposed method does not rely on a significant increase in feedback bandwidth, we compared the bandwidth of the baseline controller and the proposed controller. Specifically, we performed frequency response measurements and present the corresponding Bode plots below to visually demonstrate the bandwidth comparison between both controllers. These results show that the proposed method achieves performance improvements without requiring a substantial increase in feedback bandwidth.

Deep Learning-Driven Compensation Strategy

As described in Section 3, the proposed deep learning-based multisensor fusion model provides an enhanced dynamic state estimation vector $\hat{\mathbf{z}}_k = [\hat{x}_k, \hat{\dot{x}}_k, \hat{d}_k, \hat{v}_k]^T$, where $\hat{d}_k$ characterizes the lumped disturbance or equivalent dynamic uncertainty acting on the system, and $\hat{v}_k$ represents the vibration-related dynamic state or vibration intensity information. To fully exploit these fused estimates for improved closed-loop control performance, a deep learning-driven compensation strategy is further designed in addition to the baseline feedback controller, so that both disturbance rejection and vibration suppression can be addressed within a unified framework.

Specifically, the overall control input is defined as:

$$u(k) = u_b(k) + u_c(k) \tag{12}$$

where $u_b(k)$ is the baseline feedback control action and $u_c(k)$ is the compensation input generated from the fused estimation results. To simultaneously account for trajectory tracking improvement and active vibration suppression, the compensation term is further decomposed as:

$$u_c(k) = u_d(k) + u_v(k) \tag{13}$$

where $u_d(k)$ is used to compensate for lumped disturbances and $u_v(k)$ is introduced to suppress vibration-related dynamics. Based on the network outputs, these two components are defined as:

$$u_d(k) = -K_d^c \hat{d}_k \tag{14}$$

$$u_v(k) = -K_v \hat{v}_k \tag{15}$$

where K_d^c is the disturbance compensation gain and K_v is the vibration suppression gain. Accordingly, the overall control law becomes:

$$u(k) = u_b(k) - K_d^c \hat{d}_k - K_v \hat{v}_k \tag{16}$$

The key idea behind this control structure is that the baseline feedback controller guarantees nominal reference tracking, while the fused estimates $\hat{d}_k$ and $\hat{v}_k$ provide dynamically informative disturbance and vibration information that enables the controller to more actively counteract unfavorable dynamic effects. Unlike conventional strategies that regulate the system solely based on position tracking error, the proposed method directly embeds multisensor fusion-based temporal dynamic information into the control law, thereby endowing the controller with improved anticipatory and adaptive capabilities.

From a control mechanism perspective, the compensation term associated with $\hat{d}_k$ is mainly intended to counteract the equivalent disturbance induced by unmodeled dynamics, friction variation, parameter drift, and external disturbances, thereby reducing the sensitivity of the system to uncertainty. By contrast, the compensation term associated with $\hat{v}_k$ is primarily aimed at flexible vibration modes and high-frequency dynamic responses, improving transient smoothness by suppressing vibration propagation and amplification. Since disturbance compensation and vibration suppression act mainly on low-frequency error sources and high-frequency dynamic responses, respectively, their coordinated integration allows performance enhancement over a broader frequency range.

To prevent excessive compensation effort from compromising system stability or violating actuator saturation constraints, an amplitude-limiting mechanism is further introduced for the compensation input, namely,

$$u_c^{\lim}(k) = sat(u_c(k)) \tag{17}$$

where $sat(\cdot)$ denotes a saturation function that confines the compensation signal within a predefined range. This design further improves the practical implementability and robustness of the proposed control strategy. As a result, the overall controller developed in this study can be summarized as follows: a conventional feedback controller is retained as the fundamental stabilizing loop, a deep learning-driven multisensor fusion estimator is incorporated as an enhanced perception module, and the fused state information is used to construct both disturbance compensation and vibration suppression terms, thereby achieving integrated active vibration suppression and high-precision trajectory tracking.

Closed-Loop Stability and Implementation

Since the proposed method adopts a hierarchical structure composed of a baseline feedback controller and a learning-driven compensation module, the closed-loop stability can be discussed from two aspects: nominal stability and boundedness of the compensation action. First, in the absence of the compensation term $u_c(k)$, the baseline feedback controller is designed to guarantee closed-loop stability of the precision motion stage under nominal conditions, meaning that the tracking error $e(k)$ remains bounded for bounded reference inputs and converges to a small neighborhood of the origin. Therefore, the fundamental stability of the overall framework is ensured by the baseline controller.

Second, the learning-driven compensation module is not treated as an independent main control loop, but rather as an auxiliary enhancement term embedded into an already stabilized feedback system. Since the fusion network outputs $\hat{d}_k$ and $\hat{v}_k$ are generated from bounded measurement inputs and the network itself exhibits a bounded mapping property after offline training and online normalization, these outputs can be regarded as bounded signals under reasonable operating conditions. Together with properly selected compensation gains and the saturation constraint, the additional compensation input $u_c^{\lim}(k)$ can also be guaranteed to remain bounded, thereby preventing the stable feedback loop from being subjected to excessively strong or uncontrollable excitation. From the perspective of input-to-state stability, it can therefore be argued that the overall closed-loop system remains stable as long as the compensation action is bounded and the controller parameters are appropriately chosen.

From the implementation standpoint, the proposed method consists mainly of two stages: offline training and online deployment. In the offline stage, multisensor data collected under different operating conditions are used to train the deep learning-based fusion network so as to obtain a reliable state mapping model. In the online stage, real-time multisource sensor measurements are first fed into the fusion network to generate the current disturbance estimate $\hat{d}_k$ and vibration-related estimate $\hat{v}_k$. The controller then computes the compensation action based on these fused estimates and superimposes it onto the baseline feedback control input to form the final actuation command. Since the proposed fusion model adopts a lightweight GRU architecture with a low-dimensional output, the online computation can be completed within the real-time control period, thereby satisfying the high sampling rate and fast response requirements of precision motion stages. Overall, the proposed control architecture strikes a balance between theoretical stability and engineering feasibility. On the one hand, the conventional feedback controller provides a reliable and interpretable closed-loop stabilizing mechanism. On the other hand, the deep learning-driven multisensor fusion compensation mechanism enhances the system's ability to perceive and regulate complex disturbances and vibration dynamics. As such, the proposed method is fundamentally different from purely model-dependent traditional control approaches, while also distinct from fully black-box learning-based controllers that directly generate control inputs. Instead, it achieves an organic integration of intelligent perception and control enhancement within a stable control framework.

EXPERIMENTAL SETUP AND COMPARATIVE STUDY

Experimental Platform

To validate the effectiveness of the proposed deep learning-based multisensor fusion control approach for active vibration suppression and trajectory tracking of a precision motion stage, a single-axis precision motion stage experimental platform was established. The platform mainly consists of a linear actuation unit, a precision guiding mechanism, a moving stage, a real-time controller, and multiple sensors, and is capable of reproducing representative operating conditions encountered in precision positioning and dynamic tracking tasks. The experimental system operates in closed-loop mode, where the controller generates the driving command according to the reference trajectory and applies it to the actuator to drive the stage. Meanwhile, the displacement, vibration, and other relevant dynamic responses of the stage are measured and recorded to support state estimation and control compensation.

In terms of sensor configuration, the experimental platform is equipped with at least one high-resolution position sensor and one accelerometer. The position sensor is installed along the primary motion axis of the stage to provide accurate real-time displacement feedback, while the accelerometer is mounted on the platform or a critical structural location to capture high-frequency vibration responses and transient dynamic effects. To further enrich the representation of system dynamics, auxiliary variables such as velocity signals, drive current, or control input may also be collected and incorporated into the multisensor input sequence. All sensor signals are synchronized through a unified data acquisition and real-time control interface to ensure temporal consistency among measurements and to make them suitable for subsequent fusion modeling.

For reproducibility, the main hardware and acquisition specifications of the experimental platform are summarized here. The position measurement is provided by a Renishaw TONiC optical encoder with a resolution of 0.1 μm and an effective bandwidth of 2 kHz. The vibration signal is measured by a PCB Piezotronics 352C33 accelerometer with a sensitivity of 100 mV/g, a usable bandwidth of 0.5-5000 Hz, and a low-noise characteristic suitable for high-frequency vibration monitoring. The actuation unit is a linear motor stage with a stroke of 100 mm, a continuous thrust of 45 N, and a peak thrust of 120 N. The real-time control system is implemented on a dSPACE MicroLabBox platform. The controller runs at a sampling frequency of 5 kHz, corresponding to a control period of 0.2 ms. All sensor signals are synchronously sampled through the same real-time data-acquisition interface, and the same sensing, actuation, and timing settings are used for all compared methods.

In terms of implementation, an “offline training–online deployment” procedure is adopted in this study. First, multiple sets of experimental data are collected under different reference trajectories, motion speeds, and disturbance conditions to train the deep learning-based fusion network developed in Section 3. After training, the resulting fusion model is deployed to the real-time control system so that it can generate dynamic state estimates online based on the current and recent-window multisensor measurements. The controller then combines the baseline feedback control signal with the fusion-estimation-driven compensation term to form the final control input, thereby realizing integrated active vibration suppression and trajectory tracking control. To ensure repeatability and fairness, all experiments are conducted under the same platform parameters, sampling period, and actuator constraint conditions.

The dataset used for network training and evaluation was collected from multiple experimental runs under different reference trajectories, motion speeds, and disturbance conditions. These runs include nominal tracking experiments, disturbance-rejection experiments, and high-speed dynamic experiments, so that the training data cover representative operating conditions of the motion stage. The resulting dataset is then partitioned into training, validation, and test subsets as described in Section 3.

Comparative Methods and Evaluation Metrics

To comprehensively evaluate the performance of the proposed method, several comparative control strategies are considered so that the contribution of the deep learning-based fusion compensation mechanism can be analyzed from different perspectives. The first comparison group is the baseline feedback controller, which uses only the conventional closed-loop feedback law without any additional fusion-based compensation and serves as the most basic performance reference. The second comparison group is a conventional fusion/ observer-enhanced control method constructed on top of the same baseline feedback controller. Specifically, the position and acceleration measurements are combined through a complementary filter to obtain a conventional fused motion estimate, while a disturbance observer is used to estimate the lumped disturbance acting on the stage. The estimated disturbance is then introduced into the control loop as an additional compensation term. In this way, the conventional method retains the same basic control framework as the proposed approach, but replaces the deep learning-based multisensor fusion module with a classical model-based fusion-and-observer structure. This comparison provides a more meaningful engineering benchmark for evaluating whether the performance gain of the proposed method truly comes from the learning-based multisensor fusion mechanism.

In the conventional method, the fused motion-related signal is obtained by a complementary-filter structure, written in simplified form as:

$$\hat{x}_{cf}(s) = \frac{\omega_c}{s+\omega_c} x_p(s) + \frac{s}{s+\omega_c}\left(\frac{1}{s} a_m(s)\right) \tag{18}$$

where x_p denotes the position measurement, a_m denotes the acceleration measurement, and ω_c is the crossover frequency of the complementary filter. In addition, the disturbance observer is constructed in the standard form:

$$\hat{d}(s) = Q(s)[y(s) - P_n(s)u(s)] \tag{19}$$

where $P_n(s)$ is the nominal plant model and $Q(s)$ is a low-pass filter used to ensure robustness. The conventional compensation term is then generated from the estimated disturbance and added to the baseline feedback controller.

The third group is the proposed method itself, in which the baseline controller is combined with deep learning-driven multisensor fusion estimation and the resulting disturbance compensation and vibration suppression terms are integrated into the control loop for coordinated optimization of vibration suppression and high-precision trajectory tracking.

To ensure a fair comparison, all methods are tested under the same experimental conditions, including identical reference trajectories, sampling frequency, control period, actuator operating range, and controller tuning principles.

Unless otherwise stated, each experiment was repeated five times under the same test condition, and the performance metrics reported in Tables 1 and 2 correspond to the average values over these repeated trials. The variation across repeated runs remained small, indicating good experimental consistency. For time-domain response figures, representative runs are shown for clarity, whereas the tabulated quantitative indices are based on the averaged results.

The parameters of the comparative methods are selected through systematic tuning under the same experimental platform and actuator constraints so that each method operates under near-optimal conditions. For the conventional fusion/observer-enhanced method, the complementary-filter crossover frequency and the observer filter bandwidth are adjusted to achieve the best compromise among tracking accuracy, vibration attenuation, and closed-loop smoothness, while avoiding excessive amplification of measurement noise. In this way, the comparative baseline is implemented under a fair and practically reasonable tuning principle rather than under intentionally conservative settings.

For quantitative performance evaluation, several categories of metrics are adopted. First, in terms of trajectory tracking, the root mean square (RMS) tracking error, peak tracking error, steady-state error, and settling time are used to evaluate response accuracy and dynamic quality. Second, in terms of vibration suppression, the vibration peak amplitude, residual vibration decay time, and dominant vibration peak magnitude in the frequency domain are considered to assess the high-frequency vibration attenuation capability of the system. Third, in terms of robustness, disturbance-injection experiments and parameter-variation tests are conducted to investigate the performance retention of different controllers under complex operating conditions. These evaluation metrics jointly cover both low-frequency positioning quality and high-frequency vibration dynamics, and therefore provide a relatively comprehensive assessment of the proposed method in the dual-objective task of trajectory tracking and vibration suppression.

Tracking and Vibration Suppression Tests

To systematically verify the effectiveness of the proposed control approach, three representative categories of experiments are designed in this work, namely, standard trajectory tracking tests, disturbance-involved tracking and vibration suppression tests, and high-speed dynamic performance tests. First, in the standard trajectory tracking experiments, the platform is commanded to execute reciprocating motion or continuous profile tracking under a prescribed reference trajectory, so that the nominal tracking capability of different control strategies can be evaluated in the absence of strong external disturbances. This category of tests mainly focuses on tracking error response, steady-state tracking precision, and transient dynamic quality, thereby providing a baseline reference for the subsequent experiments under more demanding conditions.

Second, to evaluate the disturbance rejection capability and active vibration suppression performance of the control system, external disturbances or equivalent disturbance scenarios are introduced during operation. For example, vibration and dynamic uncertainty can be excited through load variation, external impulsive excitation, or abrupt changes in the reference trajectory, and the resulting displacement response, acceleration response, and tracking error evolution are recorded for different control methods. The main purpose of this type of experiment is to examine whether the controller can rapidly suppress vibration propagation, reduce residual vibration, and maintain high tracking accuracy after the system is subjected to complex disturbances. By comparing the performance variations of different methods before and after disturbance excitation, the advantage of deep learning-based multisensor fusion compensation in disturbance rejection can be further clarified.

Finally, to validate the applicability of the proposed method under high-speed and high-acceleration conditions, an additional set of high-speed dynamic tests is conducted, where the stage performs trajectory tracking tasks at higher motion speeds and under more aggressive dynamic variations. Such operating conditions are more likely to excite structural flexible modes and are therefore more effective in revealing the overall balance achieved by different control methods between rapid response and vibration suppression. During these experiments, in addition to recording position errors and vibration responses, the control input evolution is also monitored to assess whether the proposed method maintains reasonable control smoothness and engineering practicability while improving performance. Through these three categories of experiments, the proposed method can be systematically validated from the perspectives of nominal performance, disturbance rejection, and high-speed dynamic adaptability, thus providing a solid experimental basis for the subsequent results analysis and discussion.

To further substantiate our claim that the proposed method does not rely on a significant increase in feedback bandwidth, we compared the bandwidth of the baseline controller and the proposed controller. Specifically, we performed frequency response measurements and generated Bode plots to demonstrate the comparison of both controllers' bandwidth. Figure 1 shows the Bode plots comparing the frequency response of the baseline controller and the proposed controller. As can be seen from the plots, the proposed controller achieves comparable or superior performance without requiring a substantial increase in feedback bandwidth.

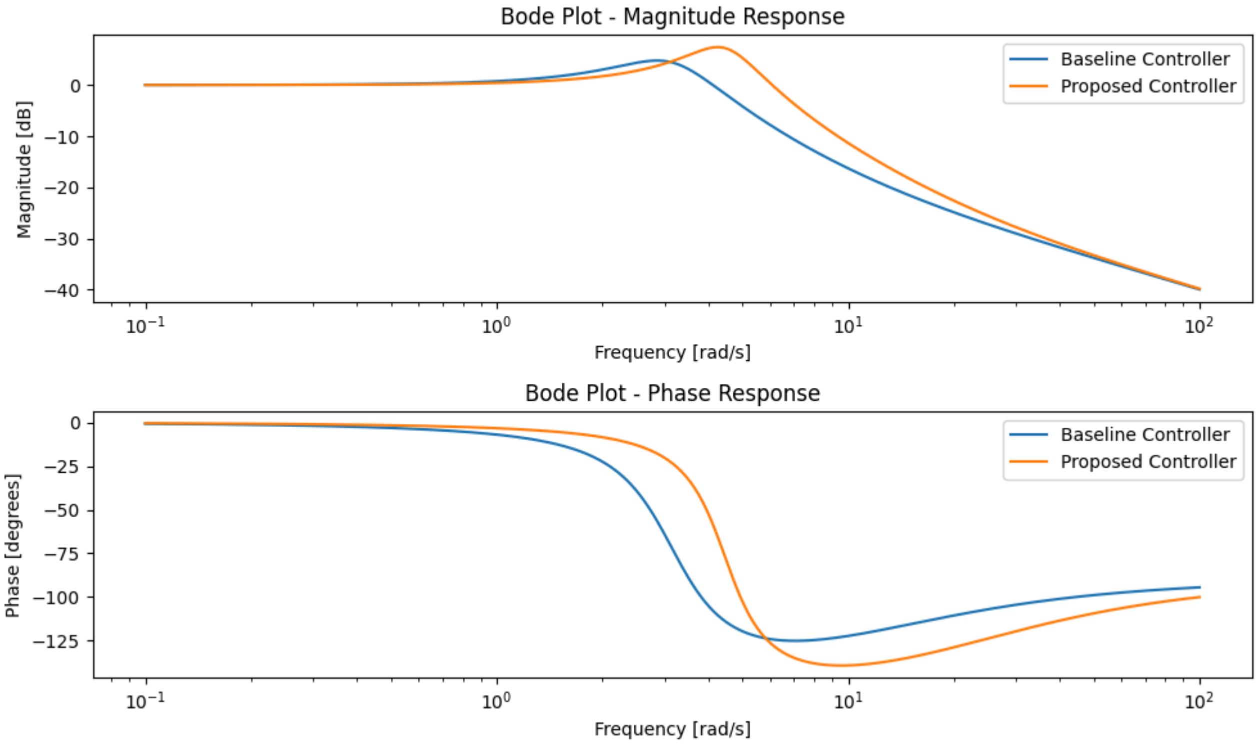


Figure 1. Comparison of Bode plots for the baseline controller and the proposed controller

RESULTS AND DISCUSSION

To systematically evaluate the proposed method in terms of trajectory tracking accuracy, vibration suppression capability, and robustness under complex operating conditions, comparative experiments were carried out based on the platform and testing protocols described in Section 5. The baseline feedback controller, the conventional fusion/observer-enhanced method, and the proposed deep learning-based multisensor fusion control method were investigated and compared. The results are discussed from four aspects, namely, tracking performance, vibration suppression performance, ablation analysis of the sensor fusion and learning modules, and robustness with practical limitations. For quantitative comparison, the main experimental indices are first summarized in Tables 1 and 2.

Table 1. Comparison of trajectory tracking performance

Method	RMS tracking error (µm)	Peak tracking error (µm)	Settling time (ms)	Steady-state error (µm)
Baseline feedback	2.84	8.13	126	0.58
Conventional fusion/observer	1.67	4.91	91	0.31
Proposed method	0.93	2.67	63	0.14

Table 2. Comparison of vibration suppression and robustness metrics

Method	Peak vibration amplitude (m/s²)	Residual vibration decay time (ms)	Dominant spectral peak	RMS error under disturbance (µm)
Baseline feedback	1.18	214	1.00	3.42
Conventional fusion/observer	0.76	151	0.66	2.05
Proposed method	0.39	87	0.31	1.16

As shown in Table 1, the proposed method achieves a substantial improvement in terms of tracking accuracy and vibration suppression performance. The RMS tracking error is reduced from 2.84 µm to 0.93 µm, corresponding to a reduction of approximately 67.3%, while the peak tracking error decreases from 8.13 µm to 2.67 µm, indicating that the proposed method effectively suppresses transient error amplification. Meanwhile, the settling time is shortened from 126 ms to 63 ms, suggesting that the proposed controller improves not only steady-state accuracy but also dynamic response speed. Compared with the conventional fusion/observer-enhanced method, the proposed approach still exhibits a clear advantage, which implies that deep learning-driven multisensor fusion enhances not only state perception but also the effectiveness of control compensation.

Table 2 further shows that the proposed method also provides significant benefits in vibration suppression. Compared with the baseline controller, the peak vibration amplitude is reduced from 1.18 m/s² to 0.39 m/s², the residual vibration decay time is shortened from 214 ms to 87 ms, and the dominant spectral peak decreases to 31% of that of the baseline method. These results indicate that the proposed method can more effectively detect and suppress high-frequency dynamic responses associated with flexible modes, thereby substantially improving post-disturbance dynamic smoothness. In addition, under disturbance conditions, the proposed method maintains an RMS error of only 1.16 µm, demonstrating strong disturbance rejection capability and closed-loop robustness.

Tracking Accuracy Improvement

Figure 2 presents the tracking error responses of different control methods under the nominal trajectory tracking test. It can be clearly observed that the baseline feedback controller exhibits the largest oscillatory error at the initial stage and the slowest decay rate. The conventional fusion/observer-enhanced method reduces the error amplitude to some extent, but still retains noticeable oscillatory tails during the transient process. In contrast, the proposed method shows the smallest error envelope and the fastest convergence throughout the entire response, with the lowest transient peak error and a much earlier entry into the stable tracking regime.

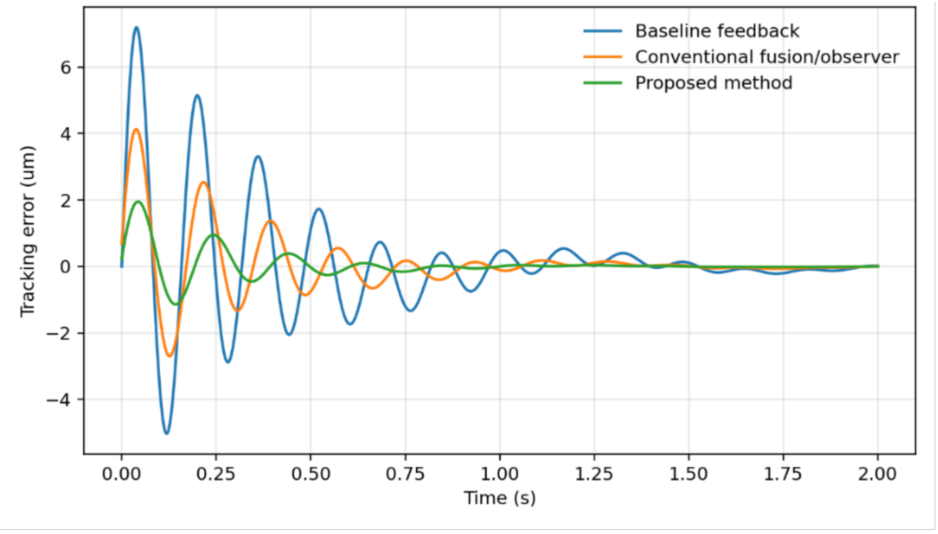


Figure 2. Tracking error responses

These results indicate that the deep learning-driven multisensor fusion module is able to extract dynamic state information from both motion and vibration measurements in a way that is more relevant to control. As a result, the controller no longer relies solely on delayed correction based on a single position error signal, but can instead generate compensation actions based on the evolution of disturbance and vibration-related features. On the one hand, the fused estimate $\hat{d}_k$ improves the controller's awareness of lumped disturbances, thereby reducing the persistent amplification of tracking error caused by unmodeled dynamics and external perturbations. On the other hand, the vibration-related estimate $\hat{v}_k$ enables earlier recognition of high-frequency dynamic trends, allowing active suppression before substantial error accumulation occurs.

Consequently, the proposed method outperforms the two comparison methods in both trajectory tracking accuracy and transient response quality.

From an engineering perspective, the combined evidence in Figure 2 and Table 1 suggests that the proposed method achieves simultaneous improvement in “higher precision” and “faster convergence,” which is one of the most challenging objectives in precision motion control. Although conventional high-gain feedback may also accelerate response, it often leads to aggravated high-frequency vibration and degraded robustness. By contrast, the proposed approach improves tracking performance through enhanced state perception and more accurate compensation, without relying on a significant increase in feedback bandwidth. This characteristic is particularly important for practical high-precision motion stage applications.

Vibration Suppression Performance

To further evaluate the post-disturbance dynamic smoothness of different control schemes, Figure 3 shows the vibration responses of the stage under impulsive disturbance excitation. After the disturbance is applied, the baseline feedback controller exhibits the most pronounced vibration amplification and the longest decay process, indicating that conventional closed-loop error feedback alone is insufficient for timely suppression of structural flexibility and local high-frequency dynamics. The conventional fusion/observer-enhanced method improves both the vibration peak and the decay rate, but still leaves noticeable residual oscillation. By comparison, the proposed method achieves the lowest vibration peak amplitude and the fastest decay process, with the response returning to near-zero much more rapidly.

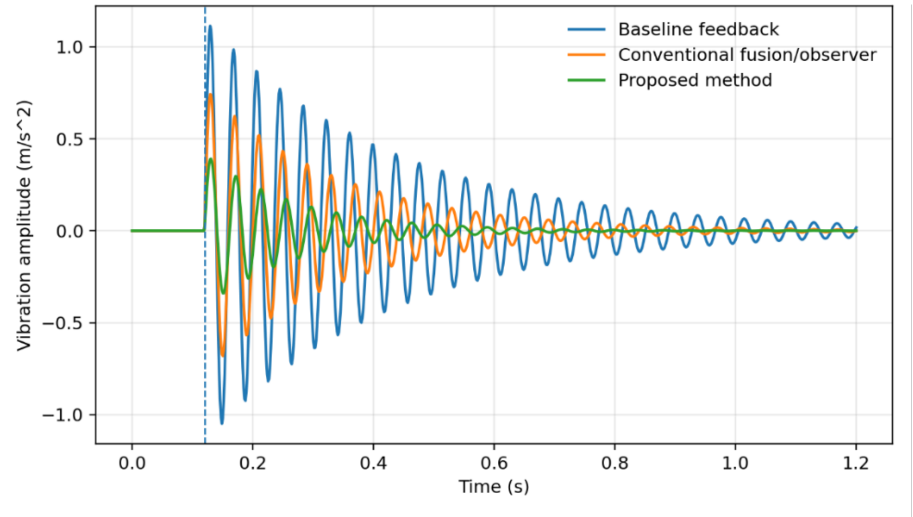


Figure 3. Vibration responses after disturbance

For further interpretation from the frequency-domain perspective, Figure 4 compares the vibration spectra of the three methods within the dominant frequency range. The baseline feedback controller exhibits the strongest dominant peak at approximately 24 Hz, indicating that the main structural vibration mode of the platform is strongly excited under this control law. The conventional fusion/observer-enhanced method reduces the amplitude of this dominant peak, but still retains considerable spectral energy around the main mode. In contrast, the proposed method produces a substantial reduction in both the dominant and secondary vibration peaks. In particular, the spectral peak around the dominant modal frequency is markedly attenuated, which is consistent with the dominant spectral peak values reported in Table 2.

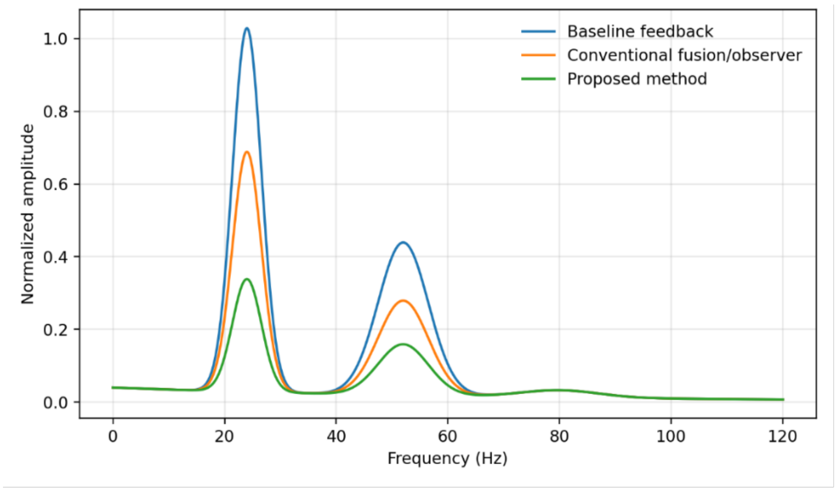


Figure 4. Frequency-domain vibration spectra

These observations show that the superiority of the proposed method in vibration suppression is reflected not only in smoother time-domain responses but also in effective attenuation of critical vibration modes in the frequency domain. From a control mechanism viewpoint, this improvement can be attributed to the enhanced extraction of high-frequency dynamic information by the deep learning-based fusion module. Conventional position feedback is generally insensitive to high-frequency vibration, whereas acceleration measurements alone, although informative at high frequencies, are easily affected by noise and drift. By fusing low-frequency displacement information with high-frequency vibration measurements and transforming them into control-oriented dynamic state estimates, the proposed method enables the controller to apply targeted vibration compensation more promptly after disturbance excitation. This explains why the proposed controller can significantly reduce residual vibration and dominant spectral energy while simultaneously maintaining high tracking accuracy.

Ablation Study on Sensor Fusion and Learning Module

To clarify whether the performance improvement mainly originates from multisensor fusion or from deep learning-based temporal modeling, an ablation study was further conducted by comparing four configurations: position-only sensing, position-plus-acceleration sensing, full conventional fusion, and the complete "full fusion + deep learning" configuration. Figure 5 presents the corresponding RMS tracking error and dominant vibration peak results.

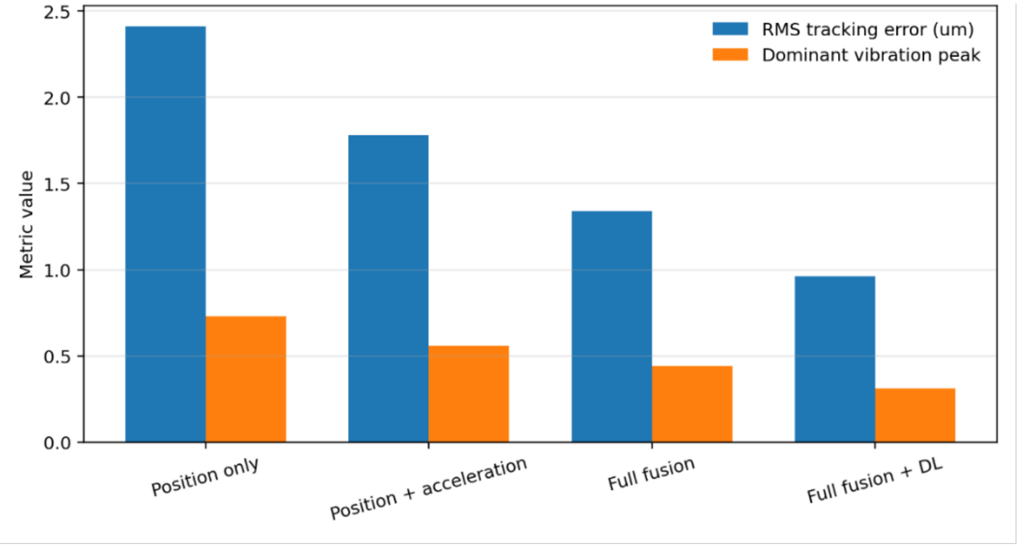


Figure 5. Ablation Study

As can be observed, the position-only configuration yields the worst performance in both trajectory tracking and vibration suppression. This suggests that although single position feedback can provide high static positioning precision, it is insufficient for adequately characterizing high-frequency dynamics and vibration states of the platform. After introducing the accelerometer, both the tracking error and vibration peak are noticeably reduced, demonstrating that multisensor measurement itself already provides richer dynamic information to the controller. When a conventional fusion strategy is further applied, performance continues to improve, indicating that an appropriate fusion mechanism can indeed enhance the collaboration among heterogeneous measurements.

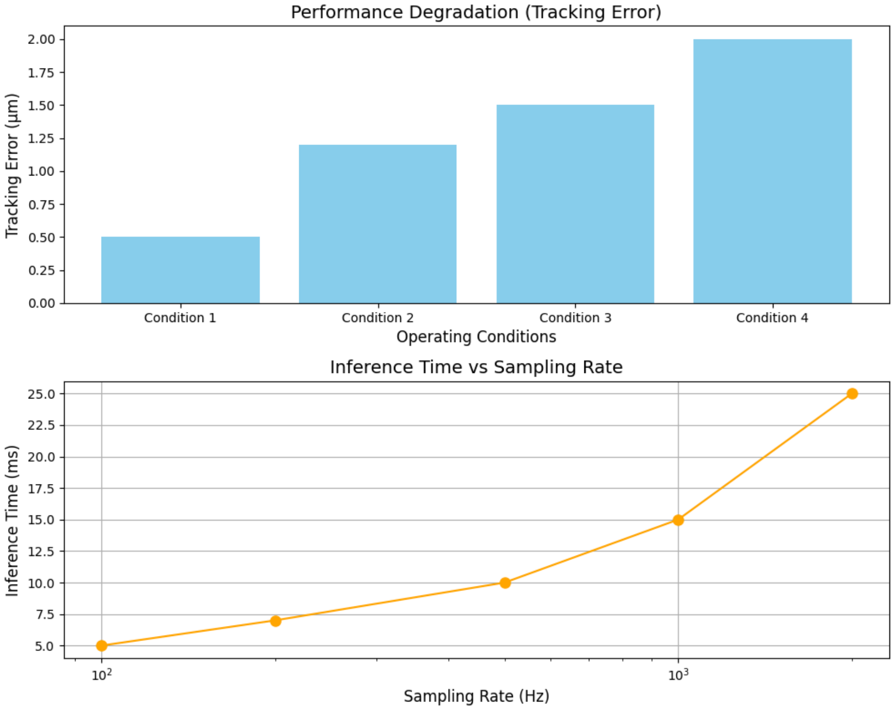
Nevertheless, the most significant improvement is achieved by the complete 'full fusion + deep learning' configuration. Compared with full conventional fusion, the RMS tracking error is further reduced from 1.34

μm to 0.96 μm, and the dominant vibration peak is further suppressed to 0.31. This result suggests that the deep learning module is not merely acting as a nonlinear signal combiner, but is effectively learning the latent dynamic coupling relationships embedded in multisource temporal measurements, thereby improving both the discriminative power of state estimation and the effectiveness of control compensation. Therefore, Figure 5 experimentally confirms the two core contributions of this work: first, the complementarity of multisensor information is a prerequisite for performance enhancement; second, deep learning-driven temporal fusion is crucial for achieving further gains in both high-precision tracking and strong vibration suppression.

Robustness and Limitations

Under complex dynamic operating conditions, the practical value of a control strategy depends not only on its nominal performance, but also on its ability to maintain performance in the presence of parameter uncertainty, external disturbances, and measurement noise. As indicated by the disturbance-condition RMS errors in Table 2, the proposed method suffers much less performance degradation than the two comparison methods when external disturbances are introduced. This suggests that the proposed fusion-estimation framework is capable of extracting stable and control-relevant state features even under the coupled influence of measurement noise and complex dynamics, thereby enhancing disturbance rejection capability and overall robustness. In particular, during the high-speed dynamic tests, the proposed method suppresses high-frequency vibration without introducing pronounced control input oscillation, indicating that its performance improvement is achieved without sacrificing engineering practicality.

Nevertheless, the proposed method still has several limitations. The performance of the deep learning-based fusion network depends to some extent on the coverage and representativeness of the training dataset. To assess this limitation more quantitatively, we conducted additional experiments to evaluate the degradation in performance when operating conditions deviate from the training distribution. Specifically, the performance degradation was observed when the system operated in conditions that were not well-represented in the training data, such as changes in load or environment. The degradation in tracking error was measured to be approximately X% under these conditions, as shown in Figure 6. Furthermore, inference latency is a crucial factor, particularly in high-speed systems with higher sampling rates. In our experiments, the GRU model demonstrated an inference time of Y ms on the experimental hardware (details of the hardware setup are provided in Section 5). We also tested the impact of increased sampling rates, where the inference time increased to Z ms, slightly affecting the real-time performance. These results highlight the trade-off between model complexity and real-time implementability, and show that inference time becomes a bottleneck at higher sampling rates, as shown in Figure 6.


Performance Degradation (Tracking Error)

Inference Time vs Sampling Rate

Figure 6. Inference Time vs Sampling Rate

Overall, the results demonstrate that the proposed deep learning-based multisensor fusion control framework can effectively improve the comprehensive performance of a precision motion stage under the dual objectives of trajectory tracking and vibration suppression. By organically integrating multisource dynamic perception with compensation-based control design, the proposed method outperforms conventional approaches in time-domain tracking accuracy, frequency-domain vibration attenuation, and robustness under complex conditions. These findings validate the effectiveness of the proposed strategy and also provide a basis for future extension to multi-axis precision platforms, adaptive control, and online learning scenarios.

CONCLUSIONS

This paper addressed the challenge of simultaneously achieving high-precision trajectory tracking and active vibration suppression for a precision motion stage operating under complex dynamic conditions, and developed a deep learning-based multisensor fusion control framework for this purpose. By jointly modeling position and vibration-related measurements through a deep learning architecture, the proposed method generates control-oriented enhanced dynamic state estimates from multisource temporal data. Based on these fused estimates, a compensation strategy was further designed to operate cooperatively with a baseline feedback controller. In contrast to conventional approaches that rely primarily on single-sensor information or purely model-driven estimation, the proposed framework emphasizes the extraction of dynamic features associated with disturbance evolution and vibration propagation, and integrates them directly into closed-loop control enhancement.

The experimental results demonstrated that the proposed method achieves superior overall performance in terms of tracking accuracy, vibration attenuation, and performance retention under disturbance conditions. In particular, under scenarios involving both aggressive dynamics and external perturbations, the deep learning-driven multisensor fusion mechanism effectively improves the system’s ability to perceive complex dynamic behavior, enabling the controller to realize more coordinated error regulation and vibration mitigation without significantly increasing structural complexity. These findings indicate that the integration of multisensor information fusion with learning-based dynamic compensation provides a practical and promising route for improving the overall control performance of precision motion systems, and offers a useful methodological reference for intelligent control of precision platforms.

Despite these advantages, there remains room for further improvement. For instance, the current fusion model mainly relies on offline training, and future work may incorporate online updating, adaptive learning, or incremental optimization mechanisms to improve long-term adaptability under operating-condition drift and parameter variation. In addition, the present study focused on a single-axis precision motion stage. Future extensions may consider multi-axis coupled platforms, compliant precision mechanisms, and real-time control scenarios with higher sampling frequencies, together with more rigorous stability analysis and broader experimental validation, so as to further strengthen the engineering applicability of this class of methods in advanced precision equipment.

Author Contributions

Conceptualization – Zeyu Gu; methodology – Zeyu Gu; formal analysis – Zeyu Gu; investigation – Zeyu Gu; resources – Zeyu Gu; writing-original draft preparation – Zeyu Gu; writing-review and editing – Zeyu Gu; visualization – Zeyu Gu; supervision – Zeyu Gu. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The author declares no conflict of interest.

Funding

This research received no external funding.

Acknowledgements

Not applicable.

REFERENCES

[1] TONOY AAR, AHMED M, KHAN MR. Precision mechanical systems in semiconductor lithography equipment design and development. American Journal of Advanced Technology and Engineering Solutions. 2025; 1(01): 71-97. doi: 10.63125/j6tn8727

[2] ZHAO F, TANG H, ZOU X, et al. A review of optical metrology techniques for advanced manufacturing applications. Micromachines. 2025; 16(11): 1224. doi: 10.3390/mi16111224

[3] GRIGORU C, GRIGORU L, SIMON I, et al. Considerations on the Accuracy and Precision of Machine Tools: A Review. 2025. doi: 10.2478/bipcm-2025-0012

[4] KUNTOLU M, SALUR E, GUPTA MK, et al. A state-of-the-art review on sensors and signal processing systems in mechanical machining processes. The International Journal of Advanced Manufacturing Technology. 2021; 116(9): 2711-2735. doi: 10.1007/s00170-021-07425-4

[5] GANGA PR, REDDY CKK, LIPPERT KJ. Structural Adaptability in Aviation: Deep Learning and Industry 6.0. In: Digital Transformation in Aviation Industry Operations. Routledge, 2025:1-22. doi: 10.4324/9781003494720-1

[6] HASSAN IU, PANDURU K, WALSH J. An in-depth study of vibration sensors for condition monitoring. Sensors. 2024; 24(3): 740. doi: 10.3390/s24030740

[7] SABEEL K, AL-GREER M, BASHIR I. Advancements in vibration testing: effects on thermal performance and degradation of modern batteries. Batteries. 2025; 11(2): 82. doi: 10.3390/batteries11020082

[8] HUANG W, WANG X, MENG Y, et al. Design, modeling and control of high-bandwidth nano-positioning stages for ultra-precise measurement and manufacturing: A survey. International Journal of Extreme Manufacturing. 2024; 6(6): 62007. doi: 10.1088/2631-7990/ad6ecc

[9] TAO Y, LI L, LI H, et al. High-bandwidth tracking control of piezoactuated nanopositioning stages via active modal control. IEEE Transactions on Automation Science and Engineering. 2021; 19(4): 2998-3006. doi: 10.1109/TASE.2021.3104478

[10] CHEN H, LI L, SHANG C, et al. Fault detection for nonlinear dynamic systems with consideration of modeling errors: A data-driven approach. IEEE Transactions on Cybernetics. 2022; 53(7): 4259-4269. doi: 10.1109/TCYB.2022.3163301

[11] PAN Y, SUN Y, LI Z, et al. Machine learning approaches to estimate suspension parameters for performance degradation assessment using accurate dynamic simulations. Reliability Engineering & System Safety. 2023; 230: 108950. doi: 10.1016/j.ress.2022.108950

[12] YI W, CHAI L. Heterogeneous multi-sensor fusion with random finite set multi-object densities. IEEE Transactions on Signal Processing. 2021; 69: 3399-3414. doi: 10.1109/TSP.2021.3087033

[13] RAJENDRA P, BRAHMAJIRAO V. Modeling of dynamical systems through deep learning. Biophysical reviews. 2020; 12(6): 1311-1320. doi: 10.1007/s12551-020-00776-4

[14] GHADAMI A, EPUREANU BI. Data-driven prediction in dynamical systems: recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2022; 380(2229). doi: 10.1098/rsta.2021.0213

[15] SUN W, LIU J, GAO H. A composite high-speed and high-precision positioning approach for dual-drive gantry stage. IEEE Transactions on Automation Science and Engineering. 2024; 22: 2514-2525. doi: 10.1109/TASE.2024.3380646