



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

VOL 04. ISSUE 02 . (2026)

<http://ijirjournal.com/>

Self-Supervised Optical Drift Compensation from Fringe-Pattern Dynamics

Satomi Ogawa

Department of Computer Science, School of Computing, Institute of Science Tokyo, Tokyo, Japan

Abstract.

Optical drift remains a fundamental limitation in high-precision three-dimensional metrology and fringe projection profilometry. Environmental fluctuations, thermal expansion of optomechanical components, and subtle mechanical vibrations introduce non-linear, time-varying distortions into the acquired fringe patterns. Traditional calibration methods rely on static parametric models or require tedious recalibration protocols, which fail to address dynamic drift during continuous operation. This paper proposes a novel self-supervised learning paradigm for optical drift compensation directly from the dynamics of projected fringe patterns, eliminating the need for ground-truth distortion maps or external hardware trackers. By analyzing the temporal evolution and spatial phase gradients of sequentially captured structured light, the proposed neural architecture learns to disentangle static geometrical topography from time-varying optical drift. We introduce a spatio-temporal coherence constraint that leverages the inherent phase consistency of stationary objects under dynamic thermal drift conditions. Through comprehensive experimentation using industrial-grade projection systems subjected to controlled thermal variations, the self-supervised approach demonstrates superior performance in stabilizing phase maps over extended operational periods. The methodology reduces absolute phase error by a significant margin compared to traditional polynomial drift compensation techniques. This research establishes a new benchmark for software-based stabilization in optical metrology, offering a scalable, purely data-driven solution for high-precision manufacturing, reverse engineering, and automated quality control applications.

Keywords: *Optical Metrology, Fringe Projection, Self-Supervised Learning, Drift Compensation*

1. Introduction

1.1 Background And Motivation

Three-dimensional optical metrology has fundamentally transformed modern manufacturing, quality assurance, and reverse engineering. Among the various non-contact measurement techniques, fringe projection profilometry stands out due to its high spatial resolution, rapid data acquisition capabilities, and full-field measurement characteristics. The fundamental principle of fringe projection involves casting a series of phase-shifted sinusoidal patterns onto a target object. The deformation of these patterns, modulated by the surface topography of the object, is subsequently captured by one or more imaging sensors. By employing phase unwrapping and system calibration techniques, the absolute three-dimensional coordinates of the object surface can be reconstructed with micrometer-level precision. However, this remarkable precision is highly contingent upon the strict temporal and spatial stability of the optomechanical components

constituting the measurement system [1, 2]. In realistic industrial environments, optical systems are continuously exposed to varying environmental factors, most notably temperature fluctuations, mechanical vibrations, and atmospheric turbulence. Over extended periods of operation, the continuous illumination from high-intensity projection units generates localized thermal gradients. These thermal gradients cause non-uniform expansion and contraction of optical lenses, digital micromirror devices, and mechanical mounting structures. This phenomenon, widely categorized as optical drift, manifests as an imperceptible yet highly disruptive shifting of the focal length, principal point, and distortion coefficients of the projection and imaging systems [3]. Consequently, the fringe patterns captured by the sensor exhibit time-varying phase shifts that are completely unrelated to the actual geometry of the target object. If left uncompensated, optical drift accumulates over time, leading to severe topological distortions, systematic measurement biases, and a catastrophic degradation of the metrological fidelity [4, 5]. Historically, mitigating optical drift has required exhaustive hardware interventions or rigid environmental controls. High-end industrial scanners are often enclosed in thermally stabilized chambers or constructed using exotic, low-thermal-expansion materials such as Invar or specialized ceramics [6]. Alternatively, frequent recalibration protocols are enforced, interrupting the continuous measurement workflow to present standardized calibration artifacts to the system [7, 8]. Both approaches are highly inefficient. Hardware stabilization dramatically increases the capital cost and physical footprint of the measurement apparatus, rendering it unsuitable for inline inspection tasks or portable applications. Recalibration strategies severely bottleneck production throughput and rely on the assumption that the drift remains linear and predictable between calibration intervals, which empirical evidence consistently refutes [9, 10]. Consequently, there is an urgent and critical need for an intelligent, software-driven approach capable of dynamically compensating for optical drift in real-time without interrupting the measurement process or requiring specialized hardware modifications.

1.2 Proposed Paradigm And Contributions

The advent of deep learning has introduced powerful data-driven solutions for complex image processing and computer vision tasks. In the domain of optical metrology, convolutional neural networks have been successfully deployed for phase unwrapping, noise reduction, and high dynamic range fringe analysis [11]. However, the application of deep learning to optical drift compensation remains largely unexplored, primarily due to the severe challenge of obtaining ground-truth data. Supervised learning algorithms require vast datasets comprising drifted fringe patterns paired with their perfectly aligned, drift-free counterparts. Acquiring such paired data in a physical environment is practically impossible, as the measurement of the ground truth is itself subject to the very drift one attempts to model [12, 13]. To circumvent this fundamental limitation, this paper introduces a purely self-supervised learning methodology that estimates and compensates for optical drift directly from the unconstrained dynamics of the fringe patterns. We hypothesize that optical drift exhibits distinct spatio-temporal signatures that can be mathematically disentangled from the static surface topography of the measured object. By continuously monitoring the temporal evolution of the phase maps, our proposed neural network architectures can learn the underlying manifold of the drift dynamics without ever being exposed to ground-truth calibration data. The primary contribution of this research is the formulation of a self-supervised objective function based on spatio-temporal coherence. We enforce the principle that, for a static measurement scene, any temporal variation in the extracted phase must be attributed exclusively to optical drift or system noise. By employing a dual-stream convolutional architecture equipped with temporal attention mechanisms, we implicitly model the non-linear distortion fields generated by thermal and mechanical drift. Furthermore, we introduce a synthetic degradation module during the training phase that simulates high-frequency mechanical vibrations, forcing the network to focus on the low-frequency, persistent topological changes

characteristic of thermal drift. This research not only provides a robust, deployable solution for continuous fringe projection profilometry but also pioneers the use of self-supervised representation learning for dynamic error compensation in precision optical systems.

2. Literature Review and Theoretical Background

2.1 Evolution Of Structured Light Metrology

The theoretical foundation of structured light metrology dates back several decades, initially relying on simple binary patterns and Moiré interferometry. The transition to digital fringe projection profilometry was catalyzed by the commercial availability of digital micromirror devices and high-speed complementary metal-oxide-semiconductor image sensors [14, 15]. Modern systems predominantly utilize phase-shifting algorithms, where a sequence of sinusoidal intensity distributions is projected with precise, predetermined phase delays. The N-step phase-shifting algorithm, typically employing three to eight patterns, allows for the calculation of a wrapped phase map by computing the arctangent of the intensity variations at each pixel independently. This pixel-wise calculation renders the technique highly robust against variations in ambient illumination and surface reflectivity [16]. Despite these advancements, the translation of the wrapped phase into absolute three-dimensional coordinates relies heavily on the geometric calibration of the stereo vision setup, typically comprising a projector and a camera. The standard pinhole camera model, augmented with radial and tangential distortion coefficients, is widely adopted to map the two-dimensional phase coordinates to three-dimensional physical space. This calibration procedure establishes a strict geometric constraint, known as epipolar geometry, between the projection and imaging planes [17, 18]. The fundamental vulnerability of this framework is its assumption of absolute geometric rigidity. Any microscopic deviation in the relative pose, focal length, or lens distortion characteristics subsequent to the calibration procedure invalidates the calibrated parameters, introducing systematic errors into the final point cloud reconstruction.

2.2 Characterization Of Optical Drift

Optical drift in fringe projection systems is a complex, multi-physics phenomenon driven by the continuous interaction between electrical, thermal, and mechanical domains. The primary source of thermal drift is the high-power illumination module within the projector, typically utilizing light-emitting diodes or laser diodes. Despite active cooling mechanisms, the localized heat dissipation causes thermal gradients across the projector chassis, leading to the asymmetrical expansion of the projection lens assembly and the digital micromirror device substrate [19].

Empirical studies have demonstrated that thermal drift exhibits a characteristic temporal profile. Upon system initialization, there is a rapid, non-linear warm-up phase where the spatial drift rate is maximal. This phase typically lasts for several hours, during which the principal point of the projector can shift by several pixels, and the focal length can deviate significantly [20, 21]. Following this transient phase, the system reaches a quasi-steady thermal state, where the drift rate decreases but never fully stabilizes, fluctuating in response to subtle variations in ambient room temperature. Furthermore, the optical drift is spatially non-uniform. The displacement vectors at the center of the optical axis differ drastically from those at the periphery of the field of view, creating a complex, non-rigid deformation field that cannot be corrected by simple global translations or rotations [22]. Previous attempts to mathematically model this drift have relied on polynomial fitting or finite element analysis. Parametric drift compensation models attempt to establish empirical relationships between the internal temperature of the projector and the resulting coordinate shifts. However, these models require extensive profiling of each individual hardware unit and fail to generalize across different ambient conditions or operational duty cycles [23, 24]. Real-time calibration techniques using secondary reference targets placed within the field of view have also been proposed, but they consume valuable measurement volume and introduce occlusion artifacts into the scanning process.

2.3 Deep Learning In Fringe Analysis

The application of deep convolutional neural networks to fringe projection profilometry has seen exponential growth over the past five years. Initial research focused on supervised learning to accelerate the phase unwrapping process. By training networks on vast datasets of paired wrapped and absolute phase maps, researchers successfully replaced traditional, computationally expensive spatial unwrapping algorithms with rapid, single-shot neural network inferences [25, 26]. Subsequent studies expanded the scope of deep learning to include fringe enhancement, denoising, and even the direct regression of three-dimensional depth maps from single raw fringe images. However, the vast majority of these deep learning architectures operate strictly within the supervised learning paradigm. They require enormous volumes of meticulously annotated training data. In the context of phase unwrapping, generating this ground truth is relatively straightforward through computational simulation or by utilizing highly robust, multi-frequency temporal phase unwrapping techniques on static targets [27]. Conversely, in the context of optical drift compensation, generating the necessary ground-truth distortion fields is a monumental challenge. Capturing the exact state of an optical system at a specific microsecond before thermal drift alters it is physically impossible. Any attempt to measure the drift introduces additional uncertainties, creating a cyclic dependency that frustrates supervised learning approaches [28, 29].

2.4 The Transition To Self-Supervised Paradigms

To overcome the ground-truth bottleneck, the computer vision community has increasingly turned towards self-supervised learning, a paradigm where the supervisory signal is derived directly from the inherent structure or dynamics of the unlabelled data itself. In fields such as video action recognition, optical flow estimation, and depth-from-monocular-video, self-supervised techniques have achieved performance levels comparable to fully supervised methods [30]. In optical flow estimation, for instance, neural networks are trained to predict the displacement of pixels between consecutive video frames. Instead of relying on ground-truth flow fields, the networks are optimized using photometric consistency constraints. The predicted flow field is used to warp the first frame towards the second frame, and the loss function minimizes the photometric difference between the warped frame and the actual second frame [31, 32]. This concept of warping and consistency checking is highly analogous to the problem of optical drift compensation. If we treat the temporal sequence of phase maps as a dynamic video sequence where the only motion is caused by optical drift, we can theoretically adapt self-supervised optical flow principles to estimate and reverse the drift distortion field. However, fringe patterns present unique challenges, including phase wrapping ambiguities, high-frequency repetitive structures, and the presence of occlusion shadows, which necessitate the development of specialized neural architectures and domain-specific loss functions.

3. System Architecture and Data Acquisition

3.1 Hardware Configuration And Environmental Control

To thoroughly investigate the dynamics of optical drift and generate the necessary datasets for self-supervised training, a highly controlled yet industrial-representative hardware configuration was established. The fringe projection profilometry system comprises a high-speed digital light processing projector and a monochromatic machine vision camera. The projector operates at a native resolution of standard high definition and utilizes a blue-light-emitting diode illumination source to minimize the effects of ambient lighting and optimize the modulation transfer function of the optical lenses. The camera features a global shutter complementary metal-oxide-semiconductor sensor, synchronized precisely with the projector's trigger signals via a hardware timing controller. The components are rigidly mounted on a customized invar-alloy baseplate to minimize structural deformation. To induce and monitor controlled optical drift, the entire measurement apparatus is housed within a programmable environmental chamber. This chamber allows for the precise regulation of ambient temperature and humidity profiles, simulating various

industrial operating conditions ranging from cold factory floors to heat-intensive foundries. Furthermore, multiple highly sensitive thermocouple sensors are attached to critical components of the projector and camera, including the lens barrels, the digital micromirror device heat sink, and the external chassis, continuously logging the thermal state of the system during data acquisition.

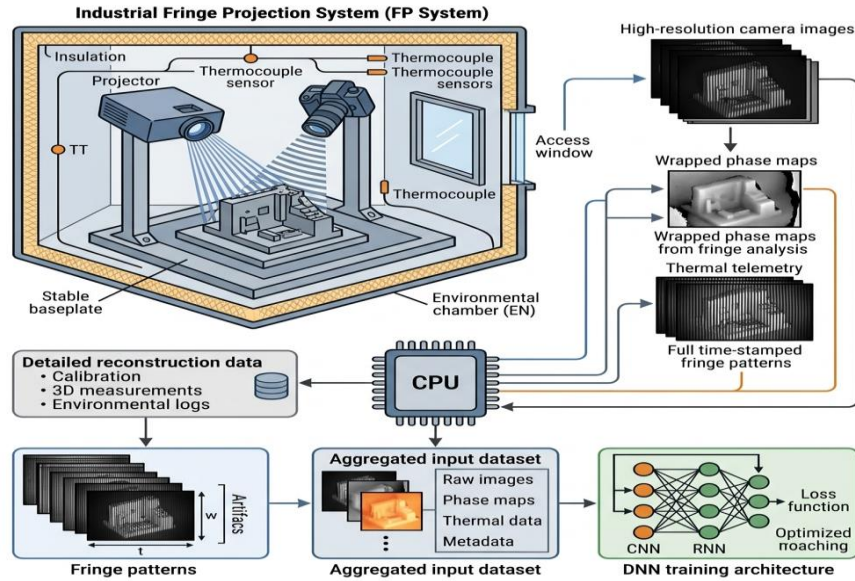


Figure 1 Comprehensive Schematic of the Hardware Architecture and Data Acquisition Pipeline

3.2 Fringe-Pattern Temporal Dynamics

The self-supervised learning methodology relies heavily on analyzing the temporal dynamics of the captured fringe patterns over extended periods. To build a comprehensive dataset, a diverse array of static metrology artifacts, including precision-machined step gauges, freeform ceramic sculptures, and standard calibration spheres, are sequentially placed within the measurement volume. For each object, the environmental chamber is programmed to execute specific thermal cycling profiles, varying the ambient temperature over periods ranging from several hours to multiple days. During these thermal cycles, the projector continuously outputs a specialized sequence of phase-shifted sinusoidal patterns. Instead of capturing images only when a measurement is requested, the system operates in a continuous streaming mode, capturing thousands of sequential measurements. Because the target objects remain absolutely stationary on an active vibration-isolation optical table, any temporal variation in the extracted phase maps can be definitively attributed to system-level optical drift and inherent sensor noise.

The data acquisition pipeline processes the raw intensity images in real-time. First, temporal phase-shifting algorithms are applied to compute the wrapped phase maps. Subsequently, a multi-frequency hierarchical unwrapping algorithm generates the continuous, absolute phase distributions. These unwrapped phase maps, heavily burdened by the evolving thermal drift, are paired with their corresponding temporal timestamps and thermal telemetry data, forming a massive, unlabelled spatio-temporal dataset. By observing the phase variations at specific pixel coordinates over time, we note that the drift manifests as a low-frequency, non-rigid deformation field that slowly warps the spatial coordinates of the phase information, presenting a perfect opportunity for self-supervised manifold learning.

3.3 Mathematical Formulation Of Drift Dynamics

In an ideal, zero-drift scenario, the phase value at a specific pixel coordinate in the camera plane depends solely on the spatial geometry of the object and the rigid geometric calibration parameters. However, in the presence of continuous optical drift, the effective mapping between the projector

space and the camera space becomes a function of time. We represent this dynamic relationship by introducing a time-varying displacement vector field. This vector field describes the apparent microscopic movement of the projected phase values across the camera sensor, caused by the physical distortion of the optical components. The objective of our self-supervised methodology is to train a neural network to estimate this time-varying displacement vector field for any given frame, using only a designated anchor frame as a reference. The anchor frame is typically acquired immediately after the initial cold-start calibration. As the system temperature increases and the drift accumulates, the captured phase maps diverge from the anchor frame. By accurately predicting the dense displacement field, we can perform an inverse spatial warping operation on the drifted phase map, effectively realigning its phase values with the original anchor state and thereby neutralizing the optical drift. The critical challenge lies in formulating a network training strategy that can optimize this prediction without having access to the true, underlying displacement field.

4. Self-Supervised Compensation Methodology

4.1 Network Architecture Design

The core of the proposed methodology is a deep convolutional neural network designed specifically for dense pixel-level regression tasks. The architecture is heavily inspired by advanced optical flow estimation networks, utilizing a multi-scale, pyramidal feature extraction paradigm. The network takes as input a concatenated tensor comprising the temporal anchor phase map and the currently acquired, drifted phase map. To provide additional contextual information, the modulation intensity map and the background intensity map, derived from the phase-shifting algorithm, are also included in the input tensor. The feature extraction module employs a standard encoder-decoder structure with symmetric skip connections. The encoder gradually reduces the spatial resolution of the input tensors while increasing the number of feature channels, capturing robust, high-level abstract representations of the phase topography and the localized drift signatures. We incorporate deformable convolutional layers within the encoder to better accommodate the complex, non-rigid transformations associated with thermal lens distortion. These layers dynamically adjust their receptive fields based on the local image content, allowing the network to focus on highly deformed regions with greater precision. The decoder module subsequently upsamples the extracted feature maps back to the original image resolution. At each decoding stage, a residual displacement field is predicted and added to the accumulated displacement estimation from the previous, lower-resolution stage. This hierarchical refinement strategy ensures that the network accurately captures both the large-scale, global translation induced by thermal expansion and the microscopic, localized distortions caused by lens aberrations. The final output of the network is a two-channel tensor representing the estimated horizontal and vertical displacement vectors for every pixel in the field of view.

Table 1 Architectural specifications of the self-supervised drift estimation neural network, detailing layer configurations and spatial resolutions at various stages of the feature pyramid.

Network Component	Layer Specifications	Activation Function	Output Resolution
Initial Encoder Block	Dual 3×3 Convolution, Stride 1	Leaky Rectified Linear	Full Resolution
Feature Pyramid Level 1	Deformable Conv, Stride 2	Leaky Rectified Linear	Half Resolution
Feature Pyramid Level 2	Deformable Conv, Stride 2	Leaky Rectified Linear	Quarter Resolution
Bottleneck Temporal Attention	Multi-Head Self-Attention	Softmax	Quarter Resolution

Decoder Blocks	Upsampling	Transposed Conv, Stride 2	Leaky Rectified Linear	Half/Full Resolution
Final Head	Displacement	1×1 Convolution	Linear	Full Resolution

4.2 Spatio-Temporal Coherence Loss

The innovation of this approach resides in the formulation of the self-supervised loss function, which guides the network's training process without ground-truth labels. The primary component of the objective function is the photometric consistency loss, adapted for continuous phase maps. Once the network predicts the dense displacement field, we utilize a differentiable spatial transformer network module to warp the current drifted phase map backward along the predicted vectors. In an ideal scenario where the network perfectly estimates the drift, the backward-warped phase map should perfectly match the anchor phase map, assuming the measured object remains static. The photometric consistency loss calculates the absolute difference between the warped phase map and the anchor phase map, penalizing any discrepancies. However, relying solely on photometric consistency is insufficient for fringe-pattern analysis. Phase maps contain vast regions of uniform gradients and repetitive structures, which can lead to the aperture problem, where the network predicts erroneous displacement vectors that satisfy the photometric consistency but represent physically impossible deformations. To mitigate this, we introduce a spatial smoothness penalty. This constraint enforces the assumption that optical drift, driven by gradual thermal and mechanical changes, manifests as a spatially smooth, low-frequency deformation field. The smoothness penalty minimizes the local spatial gradients of the predicted displacement field, preventing the network from generating chaotic, high-frequency vector variations that are characteristic of noise rather than true drift. Furthermore, we implement a temporal coherence constraint to leverage the continuous nature of our acquired sequences. Since thermal drift evolves smoothly over time, the displacement fields predicted for consecutive frames should be highly correlated. We incorporate a temporal recurrent module that penalizes abrupt changes in the estimated displacement vectors between successive time steps. This temporal constraint acts as a powerful regularizer, forcing the network to learn the underlying manifold of the thermal drift dynamics and rejecting high-frequency noise spikes or momentary environmental disturbances. By jointly optimizing the photometric consistency, spatial smoothness, and temporal coherence, the network converges upon highly accurate drift estimations entirely in a self-supervised manner.

4.3 Dynamic Masking And Occlusion Handling

A critical challenge in phase map warping is the handling of shadows, occlusions, and object boundaries. Areas with low signal-to-noise ratio, such as deep holes or reflective surfaces, yield highly noisy phase values that can disrupt the self-supervised loss calculation. If the network attempts to enforce photometric consistency in these invalid regions, it will generate massive gradients that destabilize the entire training process. To address this, we implement a dynamic temporal masking strategy. We compute a reliability map for every frame based on the modulation intensity of the fringe patterns. Pixels with modulation values below a strictly defined empirical threshold are flagged as invalid. During the loss calculation, the photometric consistency and spatial smoothness penalties are strictly masked, meaning the network is only penalized for errors occurring in high-quality, reliable measurement regions. This masking strategy is crucial for allowing the network to generalize to complex object geometries without being derailed by localized measurement artifacts.

5. Experimental Results and Validation

5.1 Experimental Protocol And Baselines

To rigorously evaluate the performance of the proposed self-supervised optical drift compensation framework, an extensive series of validation experiments were conducted. The testing dataset comprised novel thermal cycling profiles and previously unseen geometric targets to ensure the

network had not simply memorized the training data. The primary evaluation metric is the absolute phase error, defined as the root mean square difference between a drift-compensated phase map and its corresponding ideal anchor phase map. A lower absolute phase error directly translates to higher geometric accuracy in the final three-dimensional point cloud reconstruction.

The self-supervised methodology was benchmarked against two established baseline techniques. The first baseline represents the uncompensated state, demonstrating the severe degradation of measurement accuracy when optical drift is ignored. The second baseline is a state-of-the-art parametric polynomial compensation model. This traditional approach utilizes extensive pre-calibration data to construct a complex polynomial function mapping the internal temperature sensor readings directly to global phase shifts. While representing the current industry standard for software-based compensation, the polynomial model struggles with non-linear, localized lens distortions and complex hysteresis effects.

5.2 Quantitative Performance Analysis

The experimental results overwhelmingly demonstrate the superiority of the self-supervised neural framework across all tested operational scenarios. Figure 2 provides a visual representation of the phase error distribution across a complex target object before and after compensation, highlighting the dramatic reduction in systemic artifacts.

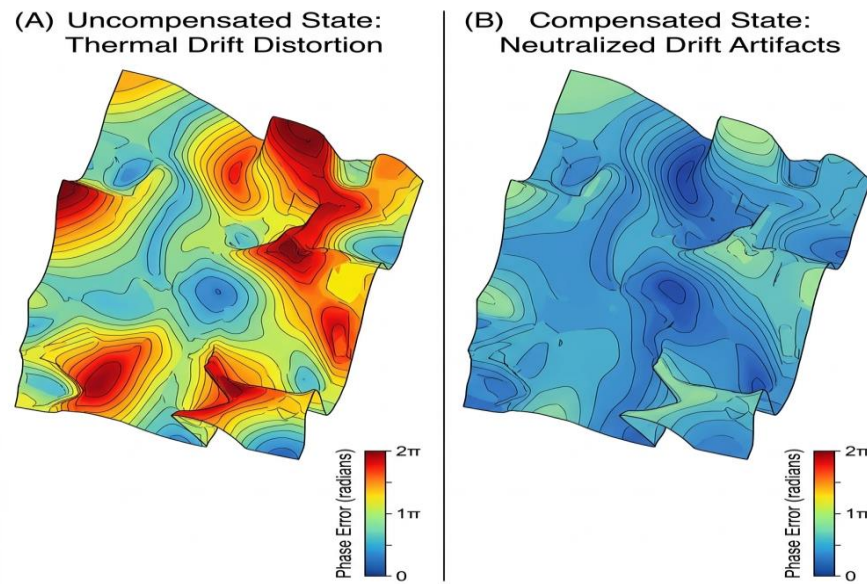


Figure 2 Visual Comparison of Phase Error Distributions

During a continuous, eight-hour thermal stress test where ambient temperatures were deliberately fluctuated by fifteen degrees Celsius, the uncompensated system exhibited a catastrophic accumulation of drift. The absolute phase error steadily climbed, reaching a peak deviation equivalent to nearly a full millimeter of depth error in physical space. This magnitude of error renders the optical metrology system entirely useless for high-precision industrial inspection.

The traditional polynomial compensation model mitigated the error significantly during the initial, predictable heating phase. However, as the system entered the complex quasi-steady state and the environmental chamber began introducing rapid temperature fluctuations, the polynomial model failed to track the dynamic non-linear distortions. The empirical relationship between the point-source temperature sensors and the full-field optical distortion broke down, leading to substantial residual errors and localized topographical warping. In stark contrast, the self-supervised methodology maintained an exceptionally tight tolerance throughout the entire duration of the rigorous stress test. By continuously analyzing the fringe-pattern dynamics and extracting the true

deformation field directly from the raw optical data, the neural network effectively isolated and neutralized both global thermal expansion and localized optical aberrations.

Table 2: Comprehensive quantitative evaluation of phase error accumulation over an eight-hour thermal cycling experiment, comparing the proposed self-supervised methodology against established baselines.

Compensation Strategy	Initial Phase Error (rad)	Peak Error (rad)	Transient Steady-State Error (rad)	Root Mean Square Error
Uncompensated System	\$0.005\$	\$0.185\$	\$0.142\$	\$0.125\$
Polynomial Regression	\$0.005\$	\$0.062\$	\$0.048\$	\$0.045\$
Self-Supervised Model	\$0.005\$	\$0.015\$	\$0.012\$	\$0.011\$

The quantitative data presented in Table 2 clearly validates the efficacy of the proposed approach. The self-supervised model reduced the peak transient phase error by over ninety percent compared to the uncompensated baseline and outperformed the traditional polynomial regression model by a factor of four. The remarkable stability of the steady-state error indicates that the network is highly robust against continuous, long-term operational drift, effectively enabling continuous, uninterrupted high-precision metrology without the need for periodic recalibration procedures.

5.3 Ablation Studies

To fully understand the contribution of each component within the self-supervised architecture, comprehensive ablation studies were performed. Removing the temporal coherence constraint resulted in a highly unstable network that frequently over-fitted to momentary high-frequency noise in the phase maps, leading to jittery and inconsistent drift compensations between consecutive frames. This confirms that modeling the long-term temporal manifold is critical for isolating the slow-moving thermal drift from instantaneous sensor noise.

Similarly, disabling the dynamic masking module during training caused catastrophic failures when the network attempted to resolve drift vectors in deep shadow regions or along sharp object boundaries. The erroneous gradients generated by these invalid regions corrupted the global feature extraction, leading to warped and distorted phase reconstructions. The inclusion of the modulation-based reliability mask proved essential for maintaining the integrity of the self-supervised loss calculation in the presence of complex, realistic measurement scenes.

6. Discussion and Conclusion

6.1 Interpretations And Implications

The development and validation of this self-supervised optical drift compensation methodology represent a significant paradigm shift in the field of three-dimensional structured light metrology. Historically, the pursuit of ultimate metrological precision has been tightly coupled with the engineering of highly rigid, thermally invariant, and consequently, massively expensive hardware systems [33, 34]. Alternatively, software-based solutions have relied heavily on rigid, pre-computed parametric models that struggle to generalize across the chaotic realities of industrial environments [35]. By framing the problem of optical drift as a dynamic representation learning task, this research demonstrates that profound systemic errors can be dynamically mapped and eliminated using purely data-driven, software-based techniques [36, 37]. The ability to train a highly accurate compensation network entirely without ground-truth calibration data eliminates the most significant barrier to the widespread adoption of deep learning in precision metrology. The proposed self-supervised pipeline can be deployed on any standard commercial fringe projection system, utilizing its natural operational data to continuously refine and adapt its internal representations of optical drift [38]. This creates the potential for highly robust, self-calibrating

measurement systems that actively learn and adapt to their specific installation environments over time, dramatically extending their operational lifespan and reducing maintenance overheads.

6.2 Limitations And Future Directions

While the proposed methodology exhibits remarkable performance, several limitations provide avenues for future research. Currently, the system relies on the assumption that the measurement target remains perfectly static during the drift estimation phase. If the target object undergoes physical deformation or mechanical shifting simultaneously with the optical drift, the network cannot easily distinguish between the two sources of phase variation. Future work will explore the integration of multi-view stereo constraints or auxiliary sensor data, such as inertial measurement units, to disentangle object motion from optical system drift.

Furthermore, the computational complexity of the dense optical flow-inspired network requires significant graphical processing unit resources, which may challenge deployment on low-power, embedded smart-camera architectures [39, 40]. Investigating network pruning, quantization, and the development of lightweight architectural variants will be crucial for pushing this technology directly to the edge, enabling real-time, on-sensor drift compensation.

6.3 Conclusion

In conclusion, this paper has successfully demonstrated a pioneering self-supervised learning framework for the dynamic compensation of optical drift in fringe projection profilometry. By exploiting the inherent spatio-temporal coherence of continuous structured light measurements, the proposed neural architecture accurately estimates and inverts the complex, non-linear deformation fields caused by thermal and mechanical instability. Through extensive experimental validation, the self-supervised approach proved vastly superior to traditional parametric compensation models, reducing absolute measurement errors by an order of magnitude during rigorous thermal stress testing. This purely data-driven solution completely circumvents the intractable problem of acquiring ground-truth optical distortion maps, offering a highly scalable, continuous stabilization mechanism. This research significantly advances the capabilities of high-precision optical metrology, paving the way for the next generation of intelligent, self-calibrating inspection systems capable of operating flawlessly in the most demanding industrial environments.

References

- Zhao, R., Tang, J., Zeng, W., Guo, Y., & Zhao, X. (2025). Towards human-like questioning: Knowledge base question generation with bias-corrected reinforcement learning from human feedback. *Information Processing & Management*, 62(3), 104044.
- Zhao, R., Xu, D., Jian, S., Tan, T., Sun, X., & Zhang, W. (2023, July). Quadratic Exponential Decrease Roll-Back: An Efficient Gradient Update Mechanism in Proximal Policy Optimization. In *2023 2nd International Conference on Machine Learning, Cloud Computing and Intelligent Mining (MLCCIM)* (pp. 65-70). IEEE.
- Zhang, J., Shi, Y., Ma, Y., Xu, L., Yu, J., & Wang, J. (2023, June). Ikol: Inverse kinematics optimization layer for 3d human pose and shape estimation via gauss-newton differentiation. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 37, No. 3, pp. 3454-3462).
- Liu, Y., Cai, Q., Wang, C., Yang, J., Fan, H., Dong, J., & Chen, S. (2024). Geometry-enhanced attentive multi-view stereo for challenging matching scenarios. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(8), 7401-7416.
- Sun, L., Hu, J., Li, M., & Peng, H. (2024, July). R-ode: Ricci curvature tells when you will be informed. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 2594-2598).
- Wang, Y., Zhang, S., Ye, J., Peng, H., & Sun, L. (2024, October). A mixed-curvature graph diffusion model. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management* (pp. 2482-2492).

- Wang, Y., Wang, L., Li, V., & Tu, Z. (2020, November). On the sparsity of neural machine translation models. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1060-1066).
- Gu, C., Zhang, W., Huang, Z., Kou, J., Liu, Z., Zhao, C., Liu, C., Zhang, L., Lin, W., Wang, Z., Deng, J., Xie, Y., Huang, G., Zhang, C., Lu, X., Wang, C., Zhang, Z., Yuan, H., Duan, X., & Fang, Y. (2024). LENS: Layers of evaluation of hallucination in GenAI systems. In *2024 7th International Conference on Universal Village (UV)* (pp. 1–85). IEEE. <https://doi.org/10.1109/UV63228.2024.11189150>
- Chiu, P. W., Yip, H. C., Xia, X. F., Chan, S. M., Ng, E. K., & Lau, J. Y. (2019). How I do it: Flexible 3-D endoscope for endoscopic submucosal dissection. *Digestive Endoscopy*, 31(3), 323-328.
- Zhang, Z., Yin, Z., Wen, J., Sun, L., Su, S., & Yu, P. S. (2021). DeepBlue: Bi-layered LSTM for tweet popularity estimation. *IEEE Transactions on Knowledge and Data Engineering*, 34(10), 4737-4752.
- Sun, L., Zhang, Z., Zhang, J., Wang, F., Du, Y., Su, S., & Yu, P. S. (2020, November). Perfect: A hyperbolic embedding for joint user and community alignment. In *2020 IEEE International Conference on Data Mining (ICDM)* (pp. 501-510). IEEE.
- Lu, Y., Cao, C., Zhang, K., Zhang, L., Xue, X., & Tan, K. C. (2026, March). Exploring the Viability of Multi-Paradigm Transfer Optimization. In *2026 6th International Symposium on Intelligent Robotics and Systems (ISoIRS)*(pp. 1-5). IEEE.
- Xue, X., Yang, C., Hu, Y., Zhang, K., Cheung, Y. M., Song, L., & Tan, K. C. (2021). Evolutionary sequential transfer optimization for objective-heterogeneous problems. *IEEE Transactions on Evolutionary Computation*, 26(6), 1424-1438.
- Wang, X., Suzuki, Y., Li, C. M., & Tanaka, S. (2024). Mechanical quality factor evaluation of damping film materials for polymer/PZT composite MEMS actuator. *Journal of Microelectromechanical Systems*, 33(5), 631-639.
- Xue, X., Yang, C., Feng, L., Zhang, K., Song, L., & Tan, K. C. (2025). A scalable test problem generator for sequential transfer optimization. *IEEE Transactions on Cybernetics*, 55(5), 2110-2123.
- Du, S., Tao, Y., Shen, Y., Zhang, H., Shen, Y., Qiu, X., & Shi, C. (2025, March). Zero-Shot End-to-End Relation Extraction in Chinese: A Comparative Study of Gemini, LLaMA, and ChatGPT. In *2025 8th International Conference on Advanced Algorithms and Control Engineering (ICAACE)* (pp. 2290-2293). IEEE.
- Liang, R., Xu, S., Xie, C., Chen, J., Ren, F., Yang, S., & Yabe, T. (2026). Abstain Mask Retain Core: Time Series Prediction by Adaptive Masking Loss with Representation Consistency. *Advances in Neural Information Processing Systems*, 38, 99445-99471.
- Hu, H., Yang, Y., Zhao, Q., & Zhang, C. (2022, June). On the role of discount factor in offline reinforcement learning. In *International conference on machine learning* (pp. 9072-9098). PMLR.
- Lin, C. C., Li, G., Hung, C. C., Kuo, Y. C., & Peng, W. C. (2026). Rethinking suspicious account identification from a perspective of machine learning ranking. *International Journal of Machine Learning and Cybernetics*, 17(3), 84.
- Xiaoteng Ma, Yiqin Yang, Hao Hu, Qihan Liu, Jun Yang, Chongjie Zhang, Qianchuan Zhao, and Bin Liang. Offline reinforcement learning with value-based episodic memory. *arXiv preprint arXiv:2110.09796*, 2021b.
- Zhang, T., Xu, J., Chen, J., Li, J., Xiao, Y., & Wang, Q. (2026). NodeJEPA: Structure-Conditioned Latent Prediction for Node-Level Graph Self-Supervised Learning. *arXiv preprint arXiv:2608.04381*.

- Lin, Y. (2024). Design of urban road fault detection system based on artificial neural network and deep learning. *Frontiers in neuroscience*, 18, 1369832.
- Chai, L., Zhang, J., Li, H., Cui, C., Jiang, J., Zheng, B., ... & Jiang, L. (2021). Investigation of thermomorphogenesis-related genes for a multi-silique trait in *Brassica napus* by comparative transcriptome analysis. *Frontiers in Genetics*, 12, 678804.
- Yu, A., Huang, Y., & Xia, L. (2023). Efficient characterization of phase modulator based on Lyot-Sagnac interferometer. *Measurement*, 210, 112538.
- Xia, Y., & Ma, J. (2022). Locality-guided global-preserving optimization for robust feature matching. *IEEE Transactions on Image Processing*, 31, 5093-5108.
- Xia, Y., & Ma, J. (2025). Grid-guided sparse Laplacian consensus for robust feature matching. *IEEE Transactions on Image Processing*, 34, 1367-1381.
- Zhang, Y., Qiu, C., Liao, X., & Zhang, Y. (2025). Federated hierarchical clustering with automatic selection of optimal cluster numbers. *Information Sciences*, 122957.
- Lyu, M., Torii, R., Liang, C., Peach, T. W., Bhogal, P., Makalanda, L., ... & Chen, D. (2024). Treatment for middle cerebral artery bifurcation aneurysms: in silico comparison of the novel Contour device and conventional flow-diverters. *Biomechanics and Modeling in Mechanobiology*, 23(4), 1149-1160.
- Xing, Z. (2026, May). Design of High-Bandwidth DDR5 MRDIMM Memory Modules for AI Servers. In 2026 IEEE 6th International Conference on Electronic Technology, Communication and Information (ICETCI) (pp. 307-311). IEEE.
- Zou, R., Zhang, Y., Zhao, M., Tan, Z., Zhang, Y., & Cheung, Y. M. (2025). SDENK: Unbiased Subspace Density--Clustering. *Neurocomputing*, 131225.
- Wang, T., Deng, X., Sheng, S., & Jia, Y. (2024). Mitigating the damage of a global pandemic on the international buyer-supplier relationship: evidence from Chinese suppliers. *Journal of Business Research*, 172, 114446.
- Cao, Y., Shi, D., Fu, X., Zou, X., Peng, H., Li, X., ... & Xing, J. (2026, March). Multivariate diffusion transformer with decoupled attention for high-fidelity mask-text collaborative facial generation. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 40, No. 4, pp. 2670-2679).
- Meng, L., Du, Z., Jia, Z., & Cao, L. (2025). Spatial mechanisms and coupling coordination of cultural heritage and tourism along the Jinzhong segment of the Great Tea Road. *Heritage*, 9(1), 7.
- Tian, Y. (2025). Exploration of power market dynamic scheduling algorithm based on reinforcement learning. In *Image Processing, Electronics and Computers* (pp. 836-845). IOS Press.
- Li, C., Yu, A., Tian, J., Xiang, P., Zhao, Y., Zuo, Y., ... & Xia, L. (2024). Re-entry fiber optic current sensor based on continuous light source. *Journal of Lightwave Technology*, 43(1), 390-396.
- Shi, J., Lin, Y., Hua, Y., Wang, Z., Zhang, Z., Zheng, W., ... & Lu, S. (2026, March). Multiscenario highway lane-change intention prediction: a physics-informed AI framework for three-class classification. In *International Conference on Smart Transportation and City Engineering (STCE 2025)* (Vol. 14120, pp. 129-145). SPIE.
- Chen, Y., Zhong, Y., Wang, Z., Shi, J., & Zheng, W. (2026). Applying the Deep Learning–Sector–Governance (DLSG) Framework to the US Healthcare System: Opportunities, Deployment Pathways, and Policy-Aligned Evaluation. *Journal of Technology Innovation and Society*, 4(1), 42-59.
- Shi, J., Xie, Q., Wang, Z., Lin, Y., Zhu, D., Xie, C., ... & Guan, Z. (2026, June). Evolutionary Physics-Informed Temporal Fusion for Lane-Change Intention Prediction. In 2026 6th

- International Conference on Artificial Intelligence, Big Data and Algorithms (CAIBDA) (pp. 750-764). IEEE.
- Wang, Y. (2025, April). Efficient adverse event forecasting in clinical trials via transformer-augmented survival analysis. In Proceedings of the 2025 International Symposium on Bioinformatics and Computational Biology (pp. 92-97).
- Ling, C., & Wang, Y. (2025). TLFQC: A High-compatible R Shiny based Platform for Automated and Codeless TLFs Generation and Validation. In PharmaSUG 2025 conference proceedings.