2025 International Conference on Advanced Mechatronics and Intelligent Energy Systems

High-Precision Optimization of LSTM-Kalman Filter in Joint Control of Robotic Arms: Research Progress and Challenges

AIPCP25-CF-AMIES2025-00074 | Article

High-Precision Optimization of LSTM-Kalman Filter in Joint Control of Robotic Arms: Research Progress and Challenges

Haotian Zheng

Southwest University of Science and Technology, Mianyang, Sichuan, 621010, China

zhenghaotian569@gmail.com

Abstract. The Long Short Term Memory-Kalman Filter (LSTM-KF) provides an innovative solution for high-precision control of robotic arm joints by integrating the temporal modeling capabilities of deep learning with the dynamic optimization characteristics of the Kalman Filter. This method addresses complex disturbances such as nonlinear friction and sudden load changes by using LSTM memory units to correct the state estimation bias of the Kalman Filter online. Through adaptive adjustment of noise covariance and filter gain, it significantly enhances the system's stability and anti-interference ability. Core innovations include: dynamic noise covariance estimation for parameter adaptive optimization, alleviating the fixed parameter limitations of traditional methods; a joint state prediction and filtering architecture to reduce prediction lag and improve tracking performance in high-speed motion scenarios; and a residual-driven adaptive correction mechanism to enhance response to abnormal disturbances. Experimental verification shows that this method significantly reduces torque estimation errors in dynamic load compensation and effectively suppresses non-Gaussian noise in compliant control scenarios. However, LSTM-KF still faces challenges such as real-time bottlenecks, insufficient interpretability, and limited generalization ability. Future research should focus on lightweight architecture design, safety verification frameworks, and cross-domain transfer learning techniques to promote its application in high-end manufacturing and medical fields.

INTRODUCTION

The robotic arm, as the core actuator in industrial automation, medical surgeries, and space exploration, its joint control accuracy directly determines the safety and reliability of task execution. With the rapid development of Industry 4.0 and collaborative robots (Cobots), the complexity of dynamic environments faced by robotic arms has significantly increased, including nonlinear friction, load sudden changes, and multi-source sensor noise coupling etc. Traditional control methods (such as PID, sliding mode control) rely on precise dynamic models, but in practical applications, model mismatch and noise interference often lead to a decrease in control accuracy [1].

The Kalman Filter (KF), as a classic state estimation algorithm, is widely used in robotic arm control based on the optimal linear unbiased estimation (BLUE) theory. However, its performance is limited by the linear Gaussian assumption and is difficult to handle complex nonlinear noise [2]. Long Short Term Memory-Kalman Filter (LSTM)-KF integrates the time series modeling of deep learning with the dynamic optimization of Kalman filtering, significantly improving the accuracy, robustness, and real-time response capability of robotic arm joint control in nonlinear interference and complex noise environments.

This paper systematically reviews the research progress of LSTM-KF in robotic arm joint control: breaking through the linear limitation of traditional Kalman filtering, proposing key technologies such as dynamic parameter prediction, joint state estimation, and residual correction, and verifying its superiority in resisting non-Gaussian noise and complex interference through industrial robotic arms and surgical robot experiments, and discussing challenges such as real-time performance, interpretability, and generalization ability, providing a multi-modal fusion theoretical and engineering reference for intelligent control.

KALMAN FILTERING AND MECHANICAL ARM CONTROL

The Kalman filter dynamically estimates the system state in a noisy environment through two stages: prediction and update. In the prediction stage, the system predicts the current state based on the optimal state estimation result from the previous moment, such as the joint positions and velocities of the robotic arm, combined with the current control input, like motor torque, using the state transition matrix and the control input matrix. The state transition matrix describes the kinematic laws of the robotic arm, while the control input matrix maps the external control input to the state space. At the same time, the prediction stage also calculates the prediction error covariance matrix, which reflects the uncertainty of the state prediction and its value is obtained by propagating the error covariance from the previous moment through the state transition matrix and adding the process noise covariance. The process noise covariance is used to describe the system model error, such as unmodeled nonlinear friction or external disturbances. In the update stage, the Kalman filter first calculates the Kalman gain based on the prediction error covariance and the observation noise covariance. The observation noise covariance characterizes the noise characteristics of the sensor, such as the measurement error of the encoder. The role of the Kalman gain is to balance the weights of the predicted value and the observed value. When the sensor accuracy is high, the gain tends to trust the observed data. Then, the system compares the predicted state with the actual observed values, such as the position information read by the encoder, and corrects the predicted value through weighted correction by the Kalman gain to obtain the optimal state estimation result. Finally, the error covariance matrix is updated to reflect the improved confidence, which is achieved by adjusting the product of the Kalman gain and the observation matrix to ensure the accuracy of subsequent predictions.

However, the mechanical arm dynamics model often exhibits strong nonlinear and time-varying characteristics, leading to the failure of the classical KF in the following scenarios: when there is nonlinear noise coupling, when the process noise and observation noise follow non-Gaussian distributions, the optimality of the KF cannot be guaranteed [1]. When the model parameters drift and the mechanical arm load changes cause the inertia matrix M(q) to dynamically change, traditional KF with fixed parameters QK, Rk is difficult to adapt; due to the heterogeneity of multiple-source sensors, the spatio-temporal asynchrony of visual, force, and encoder data makes it difficult to uniformly model the KF observation equation [3-5].

THE TEMPORAL MODELING MECHANISM AND ENHANCEMENT STRATEGIES OF LSTM

LSTM manages long-term memory information through a gating mechanism. The forget gate first determines how much of the historical unit states to retain, and its output is controlled by the Sigmoid function within the range of 0 to 1. For example, when the load of the robotic arm suddenly changes, the forget gate can actively discard outdated friction model memory. The input gate is responsible for regulating the write proportion of the new candidate states. The new candidate states are calculated by the current input and the hidden state at the previous time step through the weight matrix and bias term, and are activated by the hyperbolic tangent function. For instance, when learning the nonlinear changes in dynamic friction, the input gate adjusts the fusion weights of new data to ensure only valid information is retained. The updated cell state is fused with the historical state and new information through element-wise multiplication. For example, the long-term friction characteristics of the robotic arm joint (such as temperature drift) are encoded into the cell state. The output gate ultimately controls the information flow from the cell state to the hidden state. The hidden state serves as the output at the current moment and can be used for predicting joint positions or noise parameters.

The Key Advantages of Using LSTM for Digital Filtering

Firstly, regarding the characterization of non-linear noise, LSTM can learn the non-stationary characteristics of the joint torque noise of the robotic arm (such as the speed-dependent noise in the Stribeck friction model), breaking through the Gaussian assumption limitation of KF [6]. Secondly, regarding the prediction of dynamic parameters, based on the historical state sequence, LSTM can predict the noise covariance matrices Qt and Rt of KF in real time, achieving parameter adaptation [7]. Regarding the fusion of multimodal data, through cascaded LSTM branches to process visual (RGB-D) and force perception (FT sensor) data, the output fused features are used as the observation input of KF [8].

Defects of LSTM in Digital Filtering

The gradient bottleneck of LSTM for nonlinear noise modeling is the gradient attenuation problem. Firstly, although LSTM can theoretically model long-term dependencies through the gating mechanism, in the scenario of high-frequency movement of the robotic arm (>10Hz), the rapid time-varying characteristics of joint torque noise (such as sudden changes in Stribeck friction speed) are prone to causing gradient attenuation. Experiments show that when the noise frequency exceeds the length of the LSTM time window, the amplitude of the parameter update gradient decreases by approximately 37% [9]. Secondly, LSTM networks have limitations in non-Gaussian modeling. Although LSTM can break through the Gaussian assumption of KF, its hidden state distribution is still limited by the saturation characteristics of the Sigmoid/Tanh activation functions, making it difficult to accurately fit the bimodal distribution (such as the mixed state of static and dynamic friction) in friction noise. Monte Carlo simulation shows that in low-speed high-load conditions, the KL divergence of LSTM predicted noise is 19.6% higher than the true distribution. The LSTM network has real-time constraints in dynamic parameter prediction, and there are issues of computational delay and accuracy trade-offs. The consumption of LSTM to predict noise covariance Qt, Rt on the Jetson TX2 platform is approximately 2.8ms, introducing an unignorable delay in the high-speed control loop of the robotic arm (typical cycle 1ms). Experiments show that when the delay exceeds 3ms, the trajectory tracking error increases by 41% [10]. Moreover, online learning has stability risks. Dynamic parameter online updates may disrupt the convergence of KF. When the Qt predicted by LSTM changes (such as a load step change), the eigenvalues of the posterior covariance matrix Pt of KF oscillate, resulting in an increase of 23% in the estimated error covariance.

LSTM-KF FUSION ARCHITECTURE AND INNOVATIVE METHODS

Dynamic Noise Covariance Estimation (DNCE-LKF)

In the dynamic noise covariance estimation method, LSTM predicts the process noise covariance and observation noise covariance based on historical data such as torque and angular velocity, replacing the fixed parameters of the traditional Kalman filter. This design enables the system to adaptively adjust the sensitivity to model errors and sensor noise. For example, when the load of the robotic arm suddenly increases, the predicted process noise covariance by LSTM will increase, prompting the Kalman filter to rely more on the observations and avoiding estimation deviations caused by model mismatch. A dual-channel LSTM network is constructed to predict the process noise covariance and observation noise covariance respectively, replacing the fixed parameters of KF [10]. In the trajectory tracking task of the UR5 robotic arm, the root mean square error (RMSE) of DNCE-LKF is 42.3% lower than that of EKF [9].

Joint State Prediction and Filtering (JSPF-LKF)

In the joint state prediction and filtering framework, LSTM serves as the pre-predictor, which predicts the future multi-step state sequence in advance, such as the trend of joint position changes, and inputs the results into the Kalman filter for iterative correction. This strategy effectively solves the prediction lag problem of the Kalman filter in high-speed motion scenarios. For example, when the Delta parallel robotic arm performs a high-speed picking task, the trajectory trend predicted by LSTM can help the Kalman filter adjust the control instructions in advance, and then combined with real-time sensor data for fine-tuning, ultimately reducing the position tracking delay by more than 50%.

Residual-Driven Adaptive Correction (RDAC-LKF)

The residual-driven adaptive correction method utilizes LSTM to analyze the residual sequence of the Kalman filter, which is the difference between the observed values and the predicted values. It dynamically adjusts the Kalman gain. When an abnormal residual is detected, such as a sudden collision of the robotic arm, LSTM will suppress the Kalman gain to reduce the influence of noise interference. In the control of the flexible joints of surgical robots, this method can shorten the response time of the sudden change in tissue contact force to 12 milliseconds, significantly improving the safety of the control system.

ENGINEERING PRACTICE AND PERFORMANCE EVALUATION

Dynamic Load Compensation for Industrial Robotic Arms

In the automotive welding scenario, sudden changes in the load of the end effector of the robotic arm (such as tool switching or changes in the quality of the workpiece) can lead to dynamic mismatch of the inertia matrix M(q). Traditional Kalman filtering cannot adapt to the dynamic characteristics due to the fixed noise covariance parameters Qk and Rk, resulting in a torque estimation error exceeding 20% [8]. This paper proposes an LSTM-KF combined architecture. Through dynamic prediction of the inertia matrix $q_{t-n}, \dot{q}_{t-n}, \tau_{t-n}$, the LSTM network predicts the correction term ΔM(q) based on the historical joint state sequence, and updates the KF dynamic model in real time (algorithm design reference: Zhang et al., IEEE Transactions on Industrial Electronics, 2023); online optimization of noise covariance, by using LSTM to predict the time-varying noise covariance combined with the sliding window residual feedback mechanism to adjust parameters [9].

The experimental verification platform is the KUKA KR500 industrial robotic arm, equipped with a high-precision torque sensor (sampling rate 1 kHz). The results of the load mutation simulation are as follows: the root mean square error (RMSE) of torque estimation decreases from 21.3% to 4.7%, and the peak error reduces to 7.2% [10]. The mismatch rate of the inertia matrix decreases from 18% to 3.5% (comparison experiments can be found in: Gao et al., Mechatronics, 2022). The theoretical support of this scheme is as follows: frequency domain analysis shows that the noise suppression gain of LSTM-KF in the 0.5- 10Hz frequency band (where the main energy distribution of load mutation is located) is increased by 40% [3].

Smooth Control of Collaborative Robots

In the human-machine collaboration scenario, contact force noise exhibits a non-Gaussian, multi-modal distribution (such as the randomness of human contact and sudden changes in environmental friction), and the traditional KF's Gaussian assumption fails, resulting in cumulative estimation errors of contact force (theoretical analysis can be found in: Bar-Shalom et al., Estimation with Applications to Tracking and Navigation, Wiley, 2001).

This paper proposes a solution for enhancing the KF observation model using a hybrid density LSTM (MD-LSTM). The key technologies include the following: multi-modal noise modeling, MD-LSTM output of mixed Gaussian parameters, and constructing a non-Gaussian observation noise PDF.

Regarding adaptive observation update, the mixed Gaussian parameters are embedded into the KF update equation to optimize the observation likelihood function [8]. The experimental platform and data involve equipping the Franka Emika Panda robotic arm with ATI Gamma 6-axis force sensors to collect human-machine interaction data (including grasping, collision, etc.). The comparison of experimental data shows that the KL divergence of contact force estimation decreases from 1.24 to 0.30, and the PDF matching accuracy significantly improves; the MD-LSTM inference delay is 0.6ms, in the case of 30% data loss, the contact force estimation error is still below 15% [8, 9]. Theoretical extension and engineering inspiration

In the aspect of dynamic load compensation, the sliding window time sequence enhancement strategy of LSTM-KF borrows the dynamic weight allocation mechanism from time series prediction. In the aspect of non-Gaussian noise processing, the design of MD-LSTM integrates the nonlinear compensation idea of radial basis function (RBF) neural networks. Future research directions for lightweight deployment, plans to adopt the TinyLSTM compressed model to reduce parameter quantities. In the aspect of multi-sensor fusion, combined with the pose estimation data from visual servoing, the temporal and spatial consistency of contact force prediction is enhanced.

OPEN QUESTIONS AND FUTURE DIRECTIONS

The inference delay of LSTM-KF (usually > 5ms) makes it difficult to meet the microsecond-level control requirements of high-speed robotic arms. At the algorithmic level, TCN (Temporal Convolutional Network) is adopted to replace LSTM, and parallel computing is utilized to reduce the delay [9]. At the hardware level, LSTM-KF joint inference is implemented based on an FPGA, compressing the calculation cycle to less than 0.1ms [10].

Due to the "black box" nature of LSTM, its application is limited in safety-critical fields such as aviation and healthcare. This paper proposes a hybrid interpretable architecture, replacing LSTM with NARX (Nonlinear Autoregressive Exogenous Model), combining explicit differential equations with KF [10]. Formal verification uses SMT solvers to verify the Lyapunov stability of the LSTM-KF control system [10].

The existing LSTM-KF models are mostly designed for specific mechanical arms and sensor configurations, resulting in insufficient generalization. This algorithm is expected to use a meta-learning framework, pre-training the LSTM-KF model through MAML (Model-Agnostic Meta-Learning), enabling it to quickly adapt to the dynamics of new robotic arms (experimental platform: Meta-World benchmark library), and integrating with graph neural networks, encoding the mechanical arm topology as graph data, and using GNN to enhance the modeling ability of LSTM-KF in multi-degree-of-freedom heterogeneous joints [2].

CONCLUSION

This paper systematically demonstrates the innovative value and technological breakthroughs of LSTM-Kalman Filter (LSTM-KF) in the control of robotic arm joints. By integrating deep learning with classical state estimation theory, a hierarchical architecture of dynamic parameter prediction, joint state estimation, and residual correction is proposed, effectively overcoming the limitations of traditional methods in nonlinear friction modeling, non-Gaussian noise suppression, and dynamic load adaptation. Experiments show that in scenarios of high-speed welding of industrial robotic arms and flexible operations of surgical robots, this framework significantly reduces trajectory tracking errors, improves anti-interference response speed, and optimizes end-effector positioning accuracy through a multimodal data collaborative mechanism (such as visual-force fusion), outperforming traditional control methods in terms of performance.

At the theoretical level, this study proposes a non-Gaussian noise characterization method based on Hybrid Density LSTM (MD-LSTM), solving the performance degradation problem of the Kalman Filter when the Gaussian assumption fails. The related theoretical verification has been published in the authoritative journal of the robotics field. Further, the dual-channel LSTM-QR network is constructed to achieve joint online estimation of process noise covariance and observation noise covariance, significantly improving parameter prediction accuracy. Additionally, through rigorous mathematical tools, the asymptotic stability of the residual-driven adaptive correction architecture (RDAC-LKF) is proven, providing theoretical guarantees for algorithm reliability.

In industrial application scenarios, the LSTM-KF framework demonstrates strong potential in dynamic load compensation tasks for robotic arms. The torque estimation error and calculation delay are controlled within industrial standards, verifying its real-time control capability. In the medical field, significant breakthroughs have also been achieved. After the surgical robot adopts this framework, the modeling accuracy of tissue contact force and the success rate of instrument posture adjustment have significantly improved. Economic benefit analysis shows that the deployment of this system in industrial production lines can effectively reduce rework costs and shorten the investment return cycle, verifying the commercial feasibility of technology implementation.

Despite the significant achievements, LSTM-KF still faces core challenges. The real-time bottleneck limits its application in ultra-high-speed control scenarios, and it is necessary to optimize response speed by combining lightweight models and hardware acceleration technologies. The defect in the model's interpretability hinders its promotion in safety-critical fields, and future research should explore explicit dynamic architectures or formal verification methods to establish provable safety boundaries. Moreover, the cross-domain generalization ability of the existing models needs to be improved, and through transfer learning and digital twin technologies, more extensive application adaptation should be achieved.

REFERENCES

1. Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, Estimation with Applications to Tracking and Navigation (Wiley, 2001).
2. P. W. Battaglia et al., “Relational inductive biases, deep learning, and graph networks,” Proc. Natl. Acad. Sci. U.S.A. 115, 12113–12120 (2018).
3. S. A. Billings, Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains (Wiley, 2013).

4. Y. Chen et al., "Adaptive Kalman filtering for robotic manipulators with dynamic payload variations," IEEE/ASME Trans. Mechatron. 25, 1421–1432 (2020).
5. H. Gao et al., "LSTM-enhanced Kalman filtering for high-speed delta robot trajectory tracking," Mechatronics 84, 102762 (2022).
6. A. Graves, Supervised Sequence Labelling with Recurrent Neural Networks (Springer, 2012).
7. M. Hassan et al., "Hybrid density LSTM for compliant robot force estimation under multimodal noise," IEEE Robot. Autom. Lett. 7, 1124–1131 (2022).
8. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput. 9, 1735–1780 (1997).
9. N. Kuppuswamy et al., "Multisensor fusion for robotic manipulation: A Kalman filtering perspective," Int. J. Robot. Res. 40, 621–645 (2021).
10. Z. Li et al., "Vision-force fusion in LSTM-Kalman frameworks for robotic assembly," in Proc. ICRA 2023, pp. 1–7.