第 43 卷 第 22 期/2023 年 11 月/光学学报 研究论文

光学学报

基于注意力机制和长短期记忆网络的F-P滤波器温漂误差修正

盛文娟[1*]，胡俊[1]，彭刚定[2]

[1]上海电力大学自动化工程学院，上海 200090；

[2]新南威尔士大学电气工程与电信学院，澳大利亚 新南威尔士州 悉尼 2052

摘要 借助深度学习算法的非线性处理能力，提出基于注意力机制和长短期记忆(LSTM)网络的温漂预测模型，从而对法布里-珀罗(F-P)滤波器进行温漂补偿。针对温漂数据中复杂的时空信息，采用LSTM提取时间信息，利用注意力机制分配空间权重。实验结果表明：在升温-降温-升温环境下，所提方法和LSTM模型的最大波长漂移误差分别为6.75 pm和16.64 pm；在单调降温环境下，两种方法的最大波长漂移误差分别为5.39 pm和14.09 pm。所提方法在最大绝对误差(MAXE)、均方根误差(RMSE)和平均误差(MAE)上均优于最小二乘支持向量机(LSSVM)和循环神经网络(RNN)。

关键词 光栅；光纤光栅；法布里-珀罗滤波器；温漂误差；注意力机制；长短期记忆网络

中图分类号 O433.1 **文献标志码** A **DOI**：10.3788/AOS230879

1 引 言

可调谐法布里-珀罗(F-P)滤波器因其解调速度快、分辨率高等优点而被广泛用于光纤布拉格光栅(FBG)的波长解调[1-2]。滤波器的工作原理是在电场激励下驱动元件压电陶瓷(PZT)的逆压电效应来改变F-P滤波器的腔长[3-4]。但是，外界环境温度的变化会改变PZT的压电常数和介电常数[5-7]，进而导致F-P滤波器的解调精度降低。

为了修正F-P滤波器的漂移误差，研究人员提出了F-P标准具法[8-9]、FBG参考光栅法[10]和气体吸收谱线参考法[11]。这些方法的误差修正精度较高，但是系统成本与复杂度也较高。近年来，随着人工智能的广泛应用，采用机器学习的方法对F-P滤波器进行温漂误差修正成为一种移植性高且成本低的新选择[12]。2014年，Cheng等[13]提出一种基于粒子群优化支持向量机(SVM)的多温度变量建模方法，采用温度和温度变化率等参数作为模型的输入，补偿环形激光陀螺的温度漂移。2016年，Wang等[14]为了补偿光纤陀螺仪的温度漂移，引入了含有4个参数的组合核函数，提高了SVM的回归精度。2018年，吴军伟等[15]引入人工鱼群算法对SVM的核心参数和核函数的参数进行优化，减小了光纤陀螺的温度漂移误差。2022年，Yang等[16]采用遗传算法和反向传播神经网络算法对PZT的温漂误差进行修正。同年，Cao等[17]提出了基于温度趋势特征提取的序列样本构造方法，提高了光纤陀螺的测量精度。针对F-P滤波器的温漂问题，本课题组提出了基于集成移动窗口的最小二乘支持向量机(LSSVM)的方法对温漂误差进行修正[18]。

作为广泛使用的传统机器学习方法，SVM采用径向基函数和多项式核函数等非线性核函数来处理非线性回归问题，在小样本的非线性回归问题上表现出良好的性能。但是，SVM在建模过程并没有考虑温漂数据的时间关联信息。作为常见的深度学习方法，循环神经网络(RNN)因具有记忆功能，可从时间序列的角度建模，考虑数据样本的时间关联信息，在时间序列建模中表现良好[19]。RNN的缺点是当时间序列较长时，容易出现梯度消失或梯度爆炸的问题，导致模型难以记住长期状态[20]。

F-P滤波器的温漂数据是一个具有动态特性的时间序列，其当前的漂移量不仅取决于当前的信息，还与过去的信息密切相关。针对可调F-P滤波器在长期温度变化环境中LSSVM和RNN均对温度漂移难以准确建模的问题，本文充分考虑温漂数据中的时间和空间信息，提出基于注意力机制和长短期记忆(LSTM)网络的温漂预测模型，采用LSTM提取时间信息，利用注意力机制分配空间权重，对F-P滤波器进行温漂误差修正。

收稿日期：2023-04-26；**修回日期**：2023-05-28；**录用日期**：2023-07-03；**网络首发日期**：2023-08-01

基金项目：国家自然科学基金(61905139)、国家自然科学基金重点项目(61935002)

通信作者：*wenjuansheng@shiep.edu.cn

2　理论基础

2.1　LSTM 模型

LSTM 模型本质上是一个循环结构，但其内部结构不是一个单纯的全连接网络，而是通过门结构来控制对新旧样本的取舍，这种结构即为记忆单元。LSTM 基本模型如图 1 所示。

LSTM 模型的基本单元主要包括输入门 i_t、输出门 o_t、遗忘门 f_t。输入门的功能是向细胞状态添加新的内容，该门接受前一个单元的隐藏状态和对应时间步的输入；输出门用于显示最终的输出；遗忘门的功能是对于旧样本按比例进行丢失。该模型的数学表达式为

$$i_t=\sigma\left(W_x^{(i)}x_t+W_h^{(i)}h_{t-1}+b_i\right), \tag{1}$$

$$f_t=\sigma\left(W_x^{(f)}x_t+W_h^{(f)}h_{t-1}+b_f\right), \tag{2}$$

$$o_t=\sigma\left(W_x^{(o)}x_t+W_h^{(o)}h_{t-1}+b_o\right), \tag{3}$$

$$c_t=\varphi\left(W_x^{(c)}x_t+W_h^{(c)}h_{t-1}+b_c\right), \tag{4}$$

$$c_t=i_t\otimes c_t+f_t\otimes c_{t-1}, \tag{5}$$

$$h_t=o_t\otimes\varphi(c_t), \tag{6}$$

式中：$\otimes$ 表示 Hadamard 积运算符；i_t、f_t、o_t 分别表示输入门、遗忘门和输出门的门控单元；c_t 表示细胞状态；h_{t-1} 表示中间隐藏状态；W_x^p、W_h^p 表示不同门单元细胞状态时的状态权重参数，其中 $p\in(i,f,o,c)$；b 为偏置项；$\sigma(x)$ 为 Sigmoid 激活函数；$\varphi(x)$ 表示 tanh 激活函数。

2.2　注意力机制

注意力机制克服了传统神经网络对具有不同重要性的特征同等对待的缺点，并且直接建立输入与输出之间的映射关系，大幅加快运行速度[21]。注意力机制的模型不仅记录信息间的位置关系，还依据信息的权重来度量不同信息特征的重要性，以加强关键信息、弱化无用信息。本文采用缩放点积注意力机制，具体步骤如下：

1）将 Key 与 Query 进行点积计算得出权重。

2）使用 Softmax 函数进行归一化处理。

3）将得到的权重与对应的 Value 进行加权求和操作后得到 Attention，其计算方式为

$$\mathrm{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{Softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V}, \tag{7}$$

式中：$\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$ 分别表示 Query、Key、Value 所对应的矩阵，$\boldsymbol{Q}=\boldsymbol{K}=\boldsymbol{V}=\boldsymbol{L}$，$\boldsymbol{L}$ 为注意力层的输入；d_k 为 Query、Key、Value 的向量维度。

2.3　基于注意力机制和 LSTM 网络的温漂误差修正方法

设计了基于注意力机制和 LSTM 网络（Attention-LSTM）的温漂模型，其结构如图 2 所示。该模型主要由 LSTM 层、注意力层和线性层组成。首先，选取一定长度的时间序列样本 T，依次向后滑动，获得若干组时间序列数据，并组成多元时间训练数据集 $D(x)$。其次，采用 LSTM 算法学习当前时刻和过去时刻的输入信息，获得每个时间步的隐藏状态，将隐藏状态整合为一个上下文向量，以此作为注意力层的输入。然后，利用注意力层对数据信息进行进一步处理，通过分配权重的方式，给予重要的信息更大的权重值，突出重要信息，过滤无用信息，提高模型的预测精度[22]，在注意力层之后采用 ReLU 层来增强预测模型的非线性拟合能力[23]。最后，使用线性层进行降维，得到 F-P 滤波器的温漂预测结果。

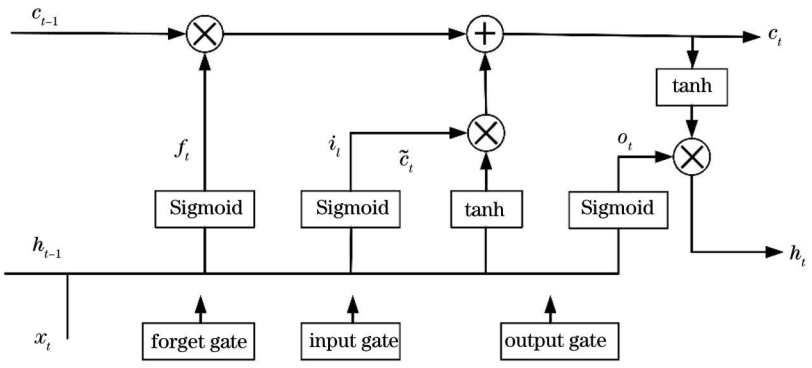

图 1　LSTM 网络结构图

Fig. 1　LSTM network structure diagram

为了评估温漂预测模型的性能，将最大绝对误差（MAXE；E_{MAX}）、均方根误差（RMSE；E_{RMS}）和平均绝对误差（E_{MA}）3 个指标作为衡量标准，对模型的预测性能进行评价。MAXE、RMSE 和 MAE 的表达式分别为

$$E_{\mathrm{MAX}}=\max\left(\left|\hat{y}_i-y_i\right|\right), \tag{8}$$

$$E_{\mathrm{RMS}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i-y_i\right)^2}, \tag{9}$$

$$E_{\mathrm{MA}}=\frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i-y_i\right|, \tag{10}$$

式中：y_i 为第 i 个真实值；$\hat{y}_i$ 为第 i 个预测值；n 为样本总数。

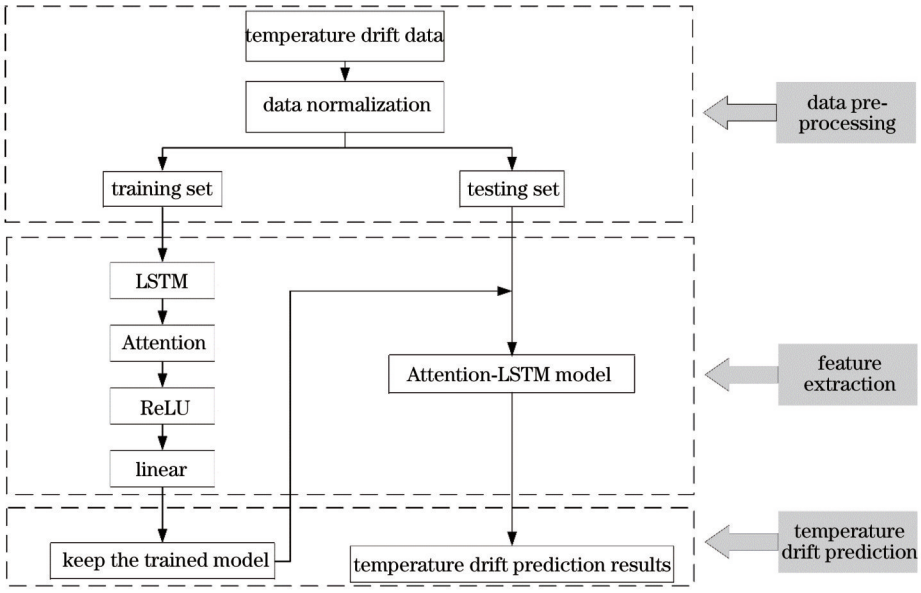

图 2　Attention-LSTM 模型的预测流程

Fig. 2　Attention-LSTM model prediction process

3　分析与讨论

3.1　实验装置

在基于可调 F-P 滤波器的 FBG 波长解调系统上进行实验，实验系统如图 3 所示。FBG 波长解调系统由光源、耦合器、FBG、F-P 滤波器、光电探测器、数据采集卡和计算机组成。光源输出宽带光，进入 4 个特征波长不同的 FBG，经 FBG 反射的光进入光电探测器，利用数据采集卡接收电压信号。此外，通过数据采集卡持续输出锯齿波电压，以驱动 F-P 滤波器工作。将 4 个 FBG 置于恒温 18 ℃的浴槽中，并将 F-P 滤波器置于温箱中，F-P 滤波器的表面贴有校准后的热敏电阻，用于维持 F-P 滤波器的表面温度。在本实验中采用 ESPEC 公司的 GSH-24V 温箱，热敏电阻温度传感器选择测温准确度为±0.001 ℃的 Fluke5641。4 个光栅的特征波长由可调激光器 HP 8164B 确定，特征波长如表 1 所示。

表 1　FBG 的特征波长

Table 1　Characteristic wavelengths of FBG

FBG	FBG0	FBG1	FBG2	FBG3
Wavelength /nm	1528.8393	1541.0621	1557.3460	1562.1832

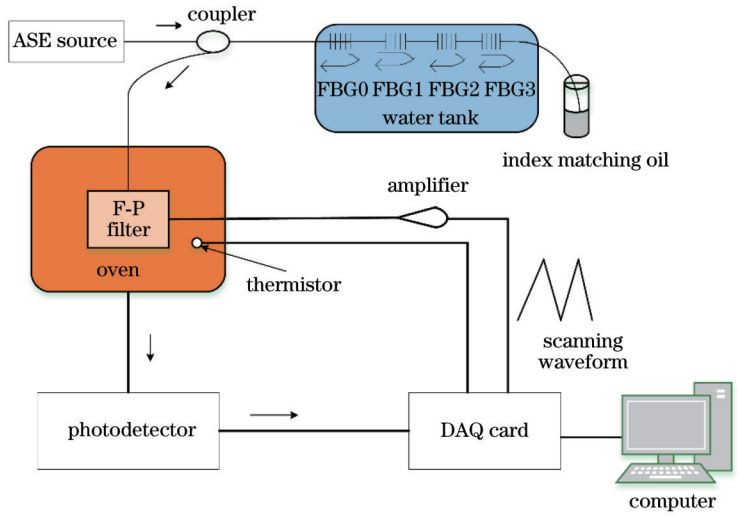

图 3　基于 F-P 滤波器的 FBG 解调系统

Fig. 3　FBG demodulation system based on F-P filter

3.2　实验与结果分析

在两种缓慢温变环境下，对所提出温漂修正模型进行实验验证。第一种温变环境下，温度从 27.5 ℃小幅升到 27.6 ℃，然后自然降至 25.6 ℃，接着升到 26.0 ℃，如图 4 所示。通过图 3 所示实验装置获取实验数据样本，将前 800 min 获取的数据作为训练数据，

后 200 min 获取的数据作为测试数据。选取温度、温度变化率和参考光栅 FBG0 的谱位置作为模型的输入特征，3 个传感光栅 FBG 的绝对波长漂移量作为模型输出。由于 F-P 滤波器的温漂结果并不仅仅由温度决定，也受到驱动电压的影响，3 个传感光栅在光谱中的位置不同，对应在锯齿波驱动电压中，扫描到该光栅的驱动电压也会不同，因此 3 个传感光栅的温漂数据是不相同的。

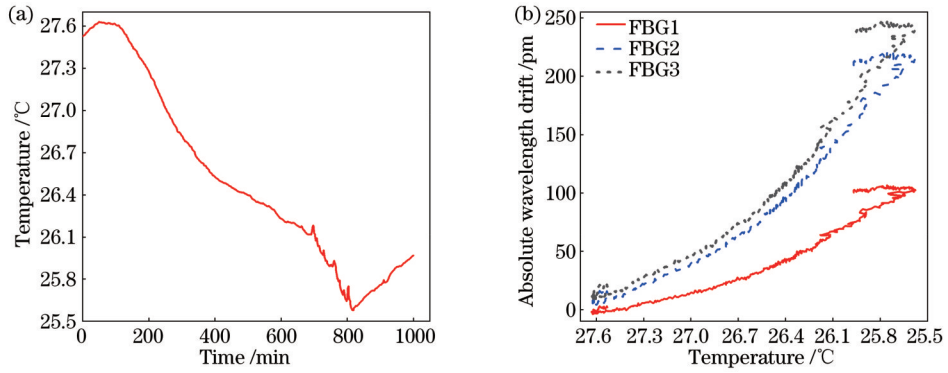

图 4　升温-降温-升温环境下 FBG 的温度测量结果与绝对波长漂移。(a)滤波器表面温度；(b)3 个传感 FBG 的绝对波长漂移

Fig. 4　Temperature measurement result and absolute wavelength drift of FBGs under heating-cooling-heating temperature variation environment. (a) Filter surface temperature; (b) absolute wavelength drift of three sensing FBGs

Attention-LSTM 模型和传统 LSTM 模型对传感光栅 FBG3 的误差修正结果如图 5 所示。传统 LSTM 模型的误差随着时间越来越大，误差呈现发散状态，最大绝对误差达到 16.64 pm。所提 Attention-LSTM 温漂模型误差收敛，最大绝对误差仅为 6.75 pm。

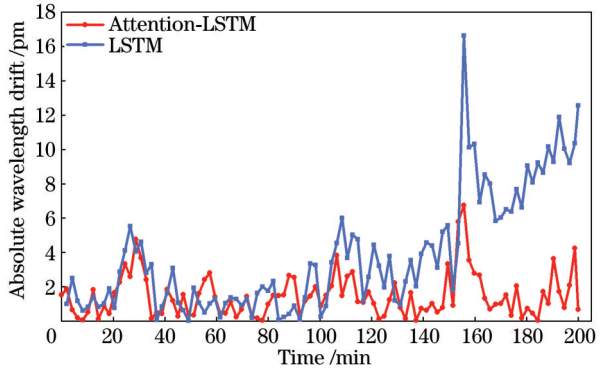

图 5　FBG3 的误差修正结果

Fig. 5　Error correction results of FBG3

Attention-LSTM 与传统 LSTM 模型对传感光栅 FBG1 和 FBG2 的误差修正结果如图 6 所示。该实验结果所得出的结论与 FBG3 的结论一致，即 LSTM 的误差随着时间的推移越来越大，模型准确度大幅降低。在测试集中，温度总体呈升温趋势，但在第 160 分钟时出现轻微的温度反复。LSTM 模型的建立依赖于前后样本的关联性，难以处理这种短暂变化，误差明显偏移；Attention-LSTM 模型不仅考虑时间关联性，还考虑输入特征的权重，因此更适应这种情况，有较好的误差回落表现。3 个 FBG 的温漂补偿模型的评价指标见表 2。对 FBG1 来说，Attention-LSTM 模型的 MAXE 比 LSTM 模型降低了 57.4%，MAE 和 RMSE 分别降低了 69.2% 和 69.7%。对 FBG2 来说，Attention-LSTM 模型的 MAXE 较 LSTM 模型降低了 58.8%，MAE 和 RMSE 分别降低了 52.4% 和 56.1%。对 FBG3 来说，Attention-LSTM 模型的 MAXE 较 LSTM 模型降低了 59.4%，MAE 和 RMSE 分别降低了 61.3% 和 62.2%。

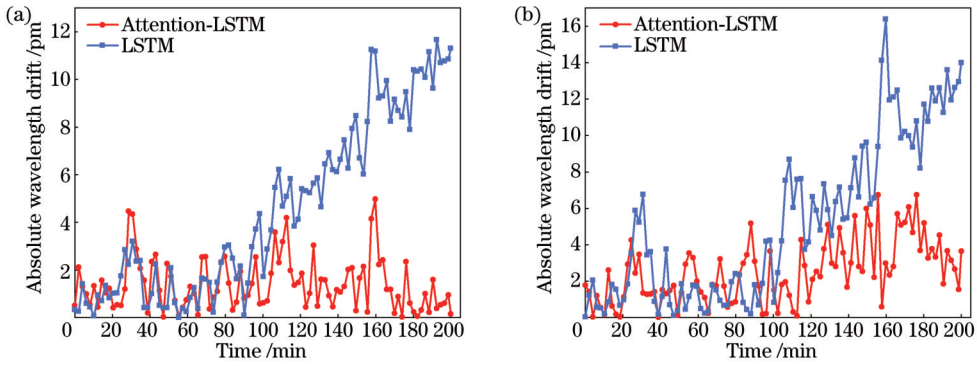

图 6　FBG1 和 FBG2 的误差修正结果。(a)FBG1；(b)FBG2

Fig. 6　Error compensation results of FBG1 and FBG2. (a) FBG1; (b) FBG2

表 2 不同FBG的温度漂移补偿误差比较

Table 2 Comparison of temperature drift compensation errors for different FBGs

FBG	FBG1		FBG2		FBG3	
	LSTM	Attention-LSTM	LSTM	Attention-LSTM	LSTM	Attention-LSTM
MAE /pm	4.52	1.39	5.25	2.50	3.77	1.46
RMSE /pm	5.81	1.76	6.85	3.01	5.10	1.93
MAXE /pm	11.67	4.97	16.39	6.75	16.64	6.75

为了进一步验证Attention-LSTM模型在不同温度变化模式下的准确建模能力，在单调降温模式下进行了实验验证。在单调降温环境下，环境温度从27 ℃自然降至25.7 ℃（即第二种降温环境），将前760 min获取的数据作为训练数据，后190 min获取的数据作为测试数据，如图7所示。由于在之前的实验中FBG3误差最大，因此选取FBG3进行实验，并且将所提模型与常见的温漂模型包括LSSVM、RNN和LSTM进行比较，实验结果如图8所示，评价指标见表3。

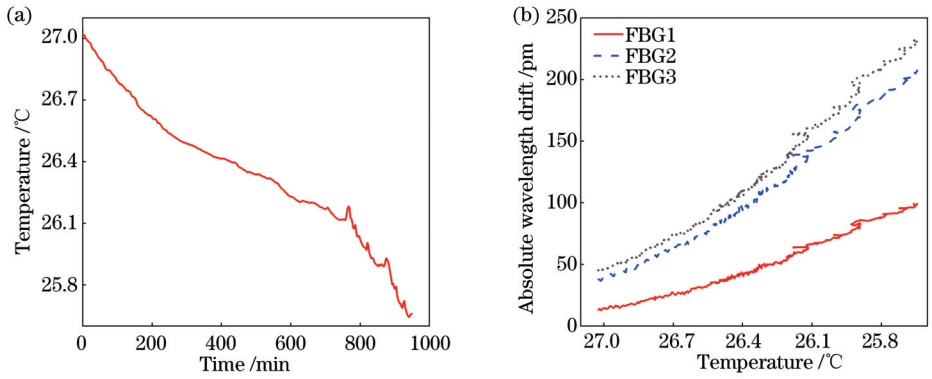

图7 单调降温环境下FBG的温度测量结果与绝对波长漂移。(a)滤波器表面温度；(b)3个传感FBG的绝对波长漂移

Fig. 7 Temperature measurement result and absolute wavelength drift of FBGs under monotone cooling environment. (a) Filter surface temperature; (b) absolute wavelength drift of three sensing FBGs

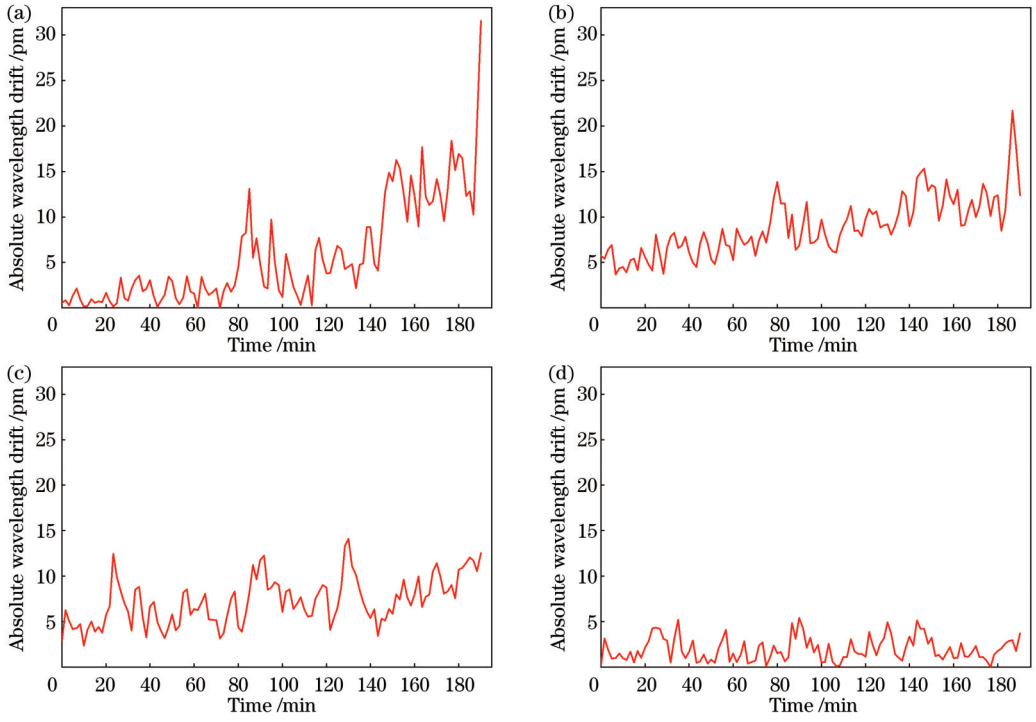

图8 单调降温环境下不同模型的波长补偿结果。(a)LSSVM；(b)RNN；(c)LSTM；(d)Attention-LSTM

Fig. 8 Wavelength compensation results of different models under monotone cooling environment. (a) LSSVM; (b) RNN; (c) LSTM; (d) Attention-LSTM

表 3 单调降温环境下各模型结果

Table 3 Results of various models under monotone cooling environment

Model	LSSVM	RNN	LSTM	Attention-LSTM
MAE /pm	5.73	8.92	4.23	2.07
RMSE /pm	8.09	9.48	7.67	2.62
MAXE /pm	31.54	21.71	14.09	5.39

实验结果表明，Attention-LSTM 模型在缓慢变化的单调降温环境下的性能优于其他模型。Attention-LSTM 与 LSSVM 模型相比，MAXE 减小了 82.9%，MAE 和 RMSE 分别减小了 63.9% 和 67.6%；与 RNN 模型相比，MAXE 减小了 75.2%，MAE 和 RMSE 分别减小了 76.8% 和 72.4%；与 LSTM 模型相比，MAXE 减小了 61.7%，MAE 和 RMSE 分别减小了 51.1% 和 65.8%。

对 4 种模型的结果进行比较，发现：LSSVM 的 MAXE 最大，这是因为建模过程没有考虑温漂数据的时间信息；RNN 从时间序列的角度建模，对短期温漂进行了记忆，故其 MAXE 比 LSSVM 有一定幅度的降低；LSTM 在 RNN 的基础上，不仅考虑了短期记忆，还考虑了长期记忆，因此其 MAXE 比 RNN 明显降低；Attention-LSTM 表现最优，因为它不仅考虑了时间上长期和短期记忆，还对特征的重要性进行了区分，为不同的输入特征分配不同的注意力权重。

4 结 论

为了对 F-P 滤波器温漂误差进行修正，在传统 LSTM 模型的基础上引入注意力机制，提出一种基于 Attention-LSTM 网络的 F-P 滤波器温漂误差修正方法。在该方法中，通过 LSTM 提取温漂数据中的时间信息，通过注意力机制为模型中的各个特征赋予不同的权重，以区分不同特征对模型的重要程度。为了验证所提算法的有效性，分别在升温-降温-升温和单调降温两种温变环境下进行了验证，并将 Attention-LSTM 模型与传统 LSTM 模型、RNN 模型和 LSSVM 模型进行比较。实验结果表明，Attention-LSTM 模型在温漂误差修正方面的性能表现明显优于其他模型，MAXE 仅为 5.39 pm，MAE 和 RMSE 分别仅为 2.07 pm 和 2.62 pm。此外，相较于传统的硬件修正方法，所提出的基于注意力机制和长短期记忆网络的 F-P 滤波器温漂误差修正方法无需引入额外的硬件设备，移植性较强、成本较低，为可调 F-P 滤波器的温漂误差修正提供了一种新的思路。

参 考 文 献

[1] Li Z Y, Xu Z Q, Tang Z H, et al. Research of high-speed FBG demodulation system for distributed dynamic monitoring of mechanical equipment[J]. Advances in Mechanical Engineering, 2013, 5: 107073.

[2] Zhao X L, Zhang Y X, Zhang W G, et al. Ultra-high sensitivity and temperature-compensated Fabry-Perot strain sensor based on tapered FBG[J]. Optics & Laser Technology, 2020, 124: 105997.

[3] Sheng W J, Peng G D, Liu Y, et al. An optimized strain demodulation method for PZT driven fiber Fabry-Perot tunable filter[J]. Optics Communications, 2015, 349: 31-35.

[4] Park H J, Song M. Linear FBG temperature sensor interrogation with Fabry-Perot ITU multi-wavelength reference [J]. Sensors, 2008, 8(10): 6769-6776.

[5] Jain A, K j P, Sharma A K, et al. Dielectric and piezoelectric properties of PVDF/PZT composites: a review[J]. Polymer Engineering & Science, 2015, 55(7): 1589-1616.

[6] Miclea C, Tanasoiu C, Amarande L, et al. Effect of temperature on the main piezoelectric parameters of a soft PZT ceramic[J]. Romanian Journal of Information Science and Technology, 2007, 10(3): 243-250.

[7] Khaliq J, Deutz D B, Frescas J A C, et al. Effect of the piezoelectric ceramic filler dielectric constant on the piezoelectric properties of PZT-epoxy composites[J]. Ceramics International, 2017, 43(2): 2774-2779.

[8] Li C, Wang Y J, Li F. Highly stable FBG wavelength demodulation system based on F-P etalon with temperature control module[J]. Infrared and Laser Engineering, 2017, 46(1): 122002.

[9] 郭海若, 刘琨, 江俊峰, 等. 基于可调谐激光器的光纤高低温力热复合多参量传感系统[J]. 中国激光, 2021, 48(19): 1906003.
Guo H R, Liu K, Jiang J F, et al. Optical fiber high and low temperature mechanical and thermal multi-parameter sensing system based on tunable laser[J]. Chinese Journal of Lasers, 2021, 48(19): 1906003.

[10] Liu K, Jing W C, Peng G D, et al. Investigation of PZT driven tunable optical filter nonlinearity using FBG optical fiber sensing system[J]. Optics Communications, 2008, 281(12): 3286-3290.

[11] 汪金辉, 许雪梅, 丁家峰, 等. 基于法布里-珀罗标准具和多光栅校准的光纤布喇格光栅波长解调系统[J]. 光子学报, 2016, 45(6): 0606003.
Wang J H, Xu X M, Ding J F, et al. Fiber Bragg grating demodulation system based on fiber Fabry-Perot etalon and multi-gratings calibrated[J]. Acta Photonica Sinica, 2016, 45(6): 0606003.

[12] 徐哲, 刘云峰, 董景新. 基于相关向量机的MEMS加速度计零偏温漂补偿[J]. 北京航空航天大学学报, 2013, 39(11): 1558-1562.
Xu Z, Liu Y F, Dong J X. Thermal bias drift compensation of MEMS accelerometer based on relevance vector machine[J]. Journal of Beijing University of Aeronautics and Astronautics, 2013, 39(11): 1558-1562.

[13] Cheng J C, Fang J C, Wu W R, et al. Temperature drift modeling and compensation of RLG based on PSO tuning SVM [J]. Measurement, 2014, 55: 246-254.

[14] Wang W, Chen X Y. Temperature drift modeling and compensation of fiber optical gyroscope based on improved support vector machine and particle swarm optimization algorithms[J]. Applied Optics, 2016, 55(23): 6243-6250.

[15] 吴军伟, 缪玲娟, 李福胜, 等. 改进支持向量机的光纤陀螺温度漂移补偿方法[J]. 红外与激光工程, 2018, 47(5): 0522003.
Wu J W, Miao L J, Li F S, et al. Compensation method of FOG temperature drift with improved support vector machine[J]. Infrared and Laser Engineering, 2018, 47(5): 0522003.

[16] Yang H L, Yang Y E, Hou Y E, et al. Investigation of the temperature compensation of piezoelectric weigh-in-motion sensors using a machine learning approach[J]. Sensors, 2022, 22 (6): 2396.

[17] Cao Y, Xu W Y, Lin B, et al. A method for temperature error

compensation in fiber-optic gyroscope based on machine learning [J]. Optik, 2022, 256: 168765.
[18] 盛文娟, 董壮志, 杨宁, 等. 基于集成移动窗口的可调谐滤波器温度补偿研究[J]. 光学学报, 2021, 41(23): 2306005.
Sheng W J, Dong Z Z, Yang N, et al. Temperature compensation of tunable filter based on integrated moving window[J]. Acta Optica Sinica, 2021, 41(23): 2306005.
[19] Pan G, Li J J, Qi Y, et al. Rapid decoding of hand gestures in electrocorticography using recurrent neural networks[J]. Frontiers in Neuroscience, 2018, 12: 555.
[20] He T, Mao H, Yi Z. Subtraction gates: another way to learn long-term dependencies in recurrent neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(4): 1740-1751.
[21] 任欢, 王旭光. 注意力机制综述[J]. 计算机应用, 2021, 41(S1): 1-6.
Ren H, Wang X G. Review of attention mechanism[J]. Journal of Computer Applications, 2021, 41(S1): 1-6.
[22] 蔡同尧, 曾献辉. 基于多特征提取的 Attention-BiGRU 短期负荷预测方法[J]. 河北电力技术, 2023, 42(1): 1-7.
Cai T Y, Zeng X H. Short term load forecasting method based on multi-feature extracted attention-BiGRU[J]. Hebei Electric Power, 2023, 42(1): 1-7.
[23] 杜启亮, 向照夷, 田联房, 等. 用于动作识别的双流自适应注意力图卷积网络[J]. 华南理工大学学报(自然科学版), 2022, 50(12): 20-29.
Du Q L, Xiang Z Y, Tian L F, et al. Double-stream adaptive attention graph convolution network for action recognition[J]. Journal of South China University of Technology (Natural Science Edition), 2022, 50(12): 20-29.

Temperature Drift Error Correction of F-P Filter Based on Attention Mechanism and LSTM Network

Sheng Wenjuan[1*], Hu Jun[1], Peng Gangding[2]

[1]*College of Automation Engineering, Shanghai University of Electric Power, Shanghai* 200090, *China*;
[2]*College of Electrical Engineering and Telecommunications, University of New South Wales, Sydney* 2052, *New South Wales, Australia*

Abstract

Objective The fiber Fabry-Perot (F-P) filter plays a critical role in fiber Bragg grating (FBG) wavelength demodulation systems. However, the continuous drift in the transmission wavelength and driving voltage curve of the F-P filter due to changes in the ambient temperature can significantly decrease the wavelength demodulation accuracy. To correct the drift error, researchers have proposed several wavelength correction methods, such as the F-P etalon, FBG reference grating, and gas absorption line reference methods. Despite high accuracy, these methods can increase the system′s cost and complexity. Recently, with the increased applications of artificial intelligence, machine learning methods have emerged as a novel and highly portable option for correcting temperature drift errors in F-P filter at a relatively low cost. Currently, the most commonly employed technique for temperature drift correction is support vector machine (SVM), which does not take into account the high temporal correlation among samples before and after temperature drift data. To this end, we propose an Attention-LSTM network-based temperature drift correction method for F-P filters. The temperature drift data for the F-P filter is a typical time series with dynamic characteristics, indicating that the current drift depends both on the present input and the past input. We adopt the LSTM model for feature extraction and apply the attention mechanism to assign different weights to various input features. The combination of short-term and long-term memory, along with the attention mechanism, enhances the demodulation accuracy of the F-P filter.

Methods We select FBG0 as the reference grating and the other three FBGs as sensing gratings. The input features employed in the model include temperature, temperature change rate, and the spectral position of FBG0. The output of the model is the absolute wavelength drift of sensing FBG3. Due to the strong temporal correlation in the temperature drift data of the F-P filter, a fixed length of time series samples is first selected, and then a multi-temporal training dataset is obtained by sliding it successively in a backward direction. By adopting the LSTM algorithm, the hidden states are generated for each time step by learning the input information of the current and past times and are then integrated into a context vector that serves as the input for the attention layer. The attention layer processes the data further by assigning weights to give significant information larger values, highlights important information, and filters out useless information, thus improving the model′s prediction accuracy. The ReLU layer is employed after the attention layer to enhance the model′s non-linear fitting abilities. Finally, a linear layer is adopted for dimensionality reduction to obtain the temperature drift prediction results of the F-P filter. The proposed model′s effectiveness is validated by comparing it to the traditional LSTM model in the same temperature environment.

Results and Discussions In a heating-cooling-heating temperature variation environment, the proposed model is compared to a traditional LSTM model in error correction results of temperature drift for the three sensing gratings (Table 2). Experimental results show that the maximum temperature drift correction error of the traditional LSTM model is 16.64 pm, while the Attention-LSTM model reduces the maximum temperature drift correction error to 6.75 pm. Additionally, the proposed model is compared to common temperature drift models such as LSSVM, RNN, and LSTM in a slowly changing monotonous cooling environment (Table 3). Experimental results indicate that the performance of the Attention-LSTM model is superior to other temperature drift models. The above experimental results demonstrate that the proposed model integrates the attention mechanism with traditional LSTM models for time series modeling. This model adopts LSTM to extract long-term and short-term data sample information over time and the attention mechanism to assign different weights to the sample features. As a result, important feature information is highlighted, and the demodulation accuracy and stability are improved.

Conclusions We thoroughly consider the time correlation between temperature drift data samples and the dynamic drift law. We not only take into account the influence of current inputs on the drift amount but also capture the effect of past inputs on the demodulation results. The LSTM model is employed for feature extraction and an attention mechanism is introduced to propose an F-P filter temperature drift error correction method based on the Attention-LSTM network. The attention mechanism assigns different weights to different features in the model to improve the modeling accuracy of the LSTM model. We conduct temperature drift error correction experiments in two temperature variation environments and compare the Attention-LSTM model with traditional LSTM model, RNN model, and LSSVM model. Experimental results demonstrate that the performance of the Attention-LSTM model in temperature drift error correction is significantly better than that of other models, with a MAXE of only 5.39 pm and MAE and RMSE of just 2.07 pm and 2.62 pm respectively. Meanwhile, compared to traditional hardware methods, the proposed error correction method based on the attention mechanism and LSTM network is low-cost with high portability, as it does not require any hardware equipment. This approach provides a new perspective for temperature drift error correction of tunable F-P filters.

Key words gratings; fiber Bragg gratings; Fabry-Perot filters; temperature drift error; attention mechanism; long short-term memory network