Reflective Microresonator based Microwave Photonic Sensor Assisted by Sparse Transformer

Xiaoyi Tian[1,2], Yeming Chen[1,2], Joel Sved[1,2], Yiming Yan[1,2], Luping Zhou, Liwei Li[1,2], Linh Nguyen, and Xiaoke Yi[1,2*]

[1]*School of Electrical and Computer Engineering, The University of Sydney, Sydney 2006, Australia*
[2]*Sydney Nano Institute, The University of Sydney, Sydney 2006, Australia*
E-mail: xiaoke.yi@sydney.edu.au

Abstract: We demonstrate a sparse transformer assisted microwave photonic sensor using a microring cascaded with an inverse designed reflector. Even with a small dataset, the root-mean-square-error of a temperature estimation model is achieved as 0.0074 °C.

1. Introduction

The microwave photonic (MWP) sensing has been receiving increasing attention due to its capability of achieving high sensing performance by transforming the measurement of optical sensor responses from the optical domain to the radio frequency (RF) domain [1]. Among different MWP sensors, those using microresonators offer advantages, including high sensitivity, small size and compact footprint [2]. Recent developments in this domain have shown that these sensors can achieve high resolution and fast speed in different applications without needing high quality (Q) factor microresonator devices, thus easing fabrication restrictions [3,4]. Moreover, by incorporating machine learning (ML) and deep learning (DL) techniques, MWP sensors with interference resistance and dual-parameter sensing capabilities have been developed [5,6]. These MWP sensors are primarily based on microresonator devices utilizing optical resonance traveling in one direction. In practice, however, due to fabrication imperfections, there exist optical resonant modes traveling in clockwise and counterclockwise directions in the same microresonator [7], with both responding to the environmental changes.

To harness the potential of both these modes for sensing, in this work, we propose a new DL assisted MWP sensing scheme using a microresonator cascaded with a reflector as the sensor probe. The reflector allows the interrogation light to pass through the microresonator twice, while the cascading effect leads to the generation of three different RF transmission notches associated with different optimal modulator bias conditions at the sensor output. By adopting sparse transformer techniques to exploit the informative features in sensor output which convey the responses of both clockwise and counterclockwise resonant modes, an improved sensing accuracy can be achieved. The proposed scheme is demonstrated in temperature sensing using a microring cascaded with an inverse designed reflector, which has a wide reflection bandwidth of over 200 nm and a tiny size of 4 μm × 4 μm. By adopting a vision transformer (ViT) neural network with modified sparse attention, a temperature estimation model is established and achieves a root-mean-square-error (RMSE) of 0.0074 °C even with a small dataset, which is over 37-fold smaller than that of the traditional linear fitting model.

2. Operation Principle

The schematic diagram of the new MWP sensor using reflective microresonator is shown in Fig. 1(a). The laser output is modulated by a frequency-swept RF signal, which is generated by a RF transceiver, in a dual-drive Mach Zehnder modulator (DDMZM) via a 90º hydride coupler. A power supply is used to adjust DDMZM bias condition. The output of the DDMZM is connected to the port 1 of an optical circulator (OC). An all-pass microring resonator

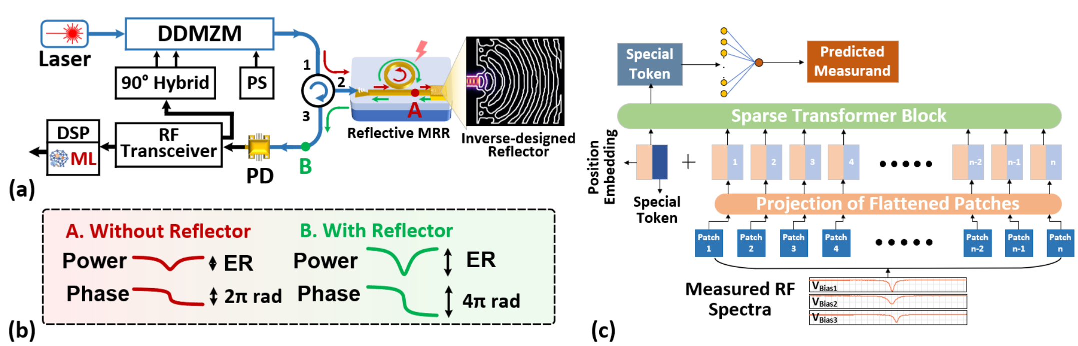

Fig. 1. (a) The schematic diagram of the proposed ML-assisted MWP sensing scheme using reflective MRR; (b) The optical power and phase transmissions of the MRR with and without a cascaded reflector; (c) The flow digram of the ViT adopted to establish the prediction model using the RF spectra acquired with three different DDMZM bias conditions. Sparse attention maps are employed to overcome overfiting. PS: power supply; DSP: digital signal processing.

(MRR) cascaded with a reflector works as an on-chip sensor probe, which has only a single optical transmission path connected to the OC port 2. The reflector is inverse designed to achieve a wide reflection bandwidth and a tiny footprint. The broadband reflector's inverse design employs the adjoint-variable method in conjunction with a level-set-based binarization approach [8, 9]. The OC port 3 is connected to the photodetector (PD). The RF transceiver measures the RF transmission after the PD and send it to a digital signal processing unit for ML based data processing.

As depicted in Fig. 1(a), as the injection light (red) transmits through the all-pass MRR from its input port to its through port, it excites the counterclockwise resonant modes. After the MRR, the light is reflected within the same waveguide path by the reflector. When the reflected light (green) travels back and transmits through the MRR the second time from its through port to its input port, it thus excites the clockwise resonant modes. In ideal situation, the counterclockwise and clockwise resonant modes have the same resonance parameters, so that the optical power and phase transmissions around the same resonance wavelength at point B is equivalent to two cascaded optical transmissions, as shown in Fig. 1(b). In practice, due to the fabrication imperfections, the resonance wavelengths of these two resonant modes often have offset, which can thus lead to dual notches and delayed phase variation in the overall optical power and phase transmissions of the reflective MRR device. According to the power and phase matching conditions between the optical sidebands and the optical resonance transmissions [3], when the selected optical resonance is over-coupled, there exist three different optimal DC bias voltages of the DDMZM to transform the overall optical resonance transmission notch obtained at point B into a deep RF transmission notch at the sensor output. To process the resulting multiple RF transmission notches at each measurement, which convey the responses of both counterclockwise and clockwise resonant modes, a ViT neural network [10] with a sparse attention module is used as shown in Fig. 1(c). In contrast to conventional neural networks, the ViT architecture excels in capturing global dependencies within signals [11]. At the first stage, the input three RF spectra are combined and divided into patches, which are then projected into a flat format. Subsequently, a special token is added at the beginning of the patch array to aggregate long-range relationship information. Position embeddings are introduced to all the patches to convey spatial information. Following this, sparse transformer blocks are employed to capture global dependencies within the input spectrum during training. After the sparse transformer blocks, the special token, representing global relationships among the three RF spectral data, is utilized to automatically map to the ground truth values using a regression neural network, achieving precise prediction of the target measurand.

3. Experimental Results

As a proof-of-concept, the proposed scheme is experimentally demonstrated for temperature sensing. Figure 2(a) and 2(b) show the reflective MRR fabricated on silicon-on-insulator platform. The reflector is designed with over 200 nm optical bandwidth and achieves a footprint of 4 µm × 4 µm. Figure 2(c) shows the optical transmission of the selected resonance of the reflective MRR measured at room temperature and compares it with that of a standard all-pass MRR with the same geometry. The reflective MRR resonance exhibits an increased extinction ratio and doubled phase change. The experimental setup is similar to that shown in Fig. 1(a). A vector network analyzer (Keysight, N5234A) is employed as the RF transceiver, which measures the RF transmission spectrum spanning from 1 GHz to 41 GHz. The sensor chip is installed on a Peltier device which is controlled by a temperature controller to adjust the chip temperature. The resonance near 1549.5 nm is selected for interrogation and the laser wavelength is fixed at 1549.33 nm. The MWP interrogation is carried out at 81 different temperatures ranging from 22.466 °C down to 20.786 °C in a step of 0.02 °C. At each temperature, three RF transmission spectra are consecutively acquired at distinct DDMZM bias voltages, each containing a notch with a rejection ratio of over 50 dB. Figure 3(a) presents the three measured RF spectra at the starting temperature. As a reference for comparison,

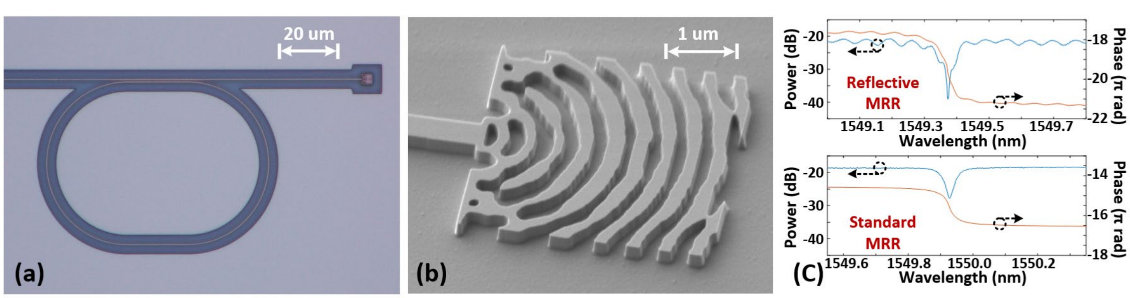

Fig. 2. (a) The microcope photo of the fabricated reflective MRR; (b) The scanning electron micrograph of the inverse-designed reflector; (c) The comparison of the optical transmissions of the reflective MRR and standard MRR.

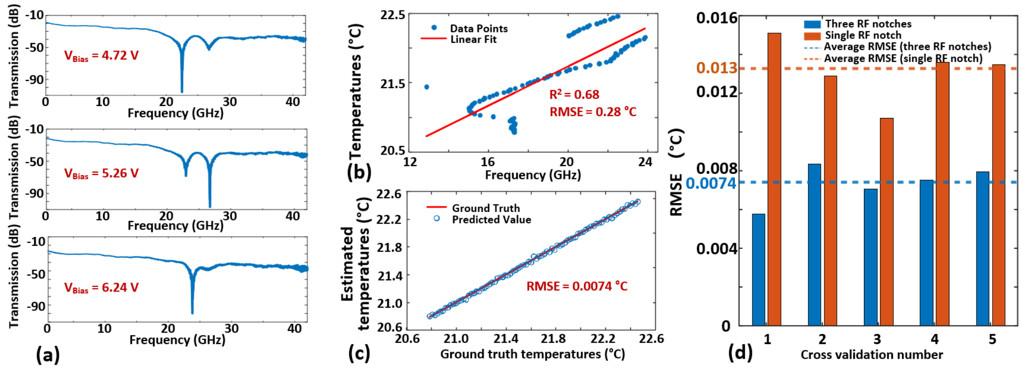

Fig. 3. (a) The measured three deep RF transmission notches at distinct DDMZM DC bias voltages (V_{Bias}); (b) The notch freqeuncy locations of the first RF spectrum in each measurement and the linear fitting results; (c) The esimated temperatures by the established ViT model against the ground truth vlaues; (d) The comparison of the RMSE performance of the ViT models esatablished by using three RF notches and using only the first RF notch.

the frequency positions of the first RF notch at each measurement are extracted and plotted along with the linear fitting results in Fig. 3(b). The R-square value of 0.68 and the RMSE of 0.28 indicate the poor performance of using this model for temperature prediction. The experimental ViT model consists of one embedding block, three attention blocks, and one multilayer perceptron (MLP) layer designed for the regression task. The embedding block partitions each combined spectra, which has a size of (3, 1000) after a max-pooling process, into 50 patches with a patch size of (3, 500), accomplished by using a stride size of (3,10). To reduce the amount of parameters of self-attention used in the original ViT, a sparse attention matrix is applied, which has 1s on the main diagonal and the two adjacent diagonals, and 0s elsewhere, considering only the correlations of three neighboring patches. This significantly enhances training efficiency and mitigates overfitting. The final MLP layer for regression uses a size of (32, 1). To further prevent overfitting, a dropout layer that randomly deactivates 5% of the nodes is introduced in both the attention map and the MLP layer, and the 5-fold cross validation is conducted. In each fold, 80% of the data is allocated for training, while the remaining 20% is reserved for testing. The learning rate is set to 0.001 and Adam optimizer is used. As shown in Fig. 3(c), the estimated temperatures by the established ViT model are notably more centralized around the ground truth values. The achieved RMSE is as low as 0.0074 °C, which is more than 37 times smaller than that of the linear fitting model in Fig. 3(b). Figure 3(d) further compares the RMSE performance of the ViT models established with three RF notch spectra and only the first RF notch spectrum in each measurement, respectively. It can be seen that utilizing all the three RF spectra, which convey the responses of both counterclockwise and clockwise resonant modes, enables 1.7 times lower RMSE in temperature prediction.

4. Conclusion

We have proposed and demonstrated a new MWP sensor using a reflective MRR. The inverse-designed reflector enables the simultaneous interrogation of both counterclockwise and clockwise resonant modes. The established sparse transformer based temperature estimation model established with the resulting three deep RF notch spectra achieves a RMSE of 0.0074 °C, which is over 1300-folder smaller than that of the traditional linear fitting model using the frequency locations of only the first RF notch.

5. Acknowledgement

The integrated devices are fabricated at Research Prototype Foundry at the University of Sydney. The authors would like to thank Smit Desai from the University of Sydney for chip preparation and measurement.

6. References

[1] J. Yao, J. Light. Technol., vol.39, no.12, pp.3626-3637, 2021.
[2] X. Jiang et al., Matter, vol.3, no.2, pp. 371-392, 2020.
[3] X. Tian et al., J. Light. Technol., vol. 38, no. 19, pp. 5440-5449, 2020.
[4] X. Tian et al., "Microwave photonic sensor based on optical sideband processing with linear frequency-modulated pulse," MWP2022, pp. 1-4.
[5] X. Tian et al., IEEE J. Sel. Top. Quantum Electron., vol. 29, no. 6, pp. 1-11, 2023.
[6] X. Tian et al., J. Light. Technol., doi: 10.1109/JLT.2023.3292261.
[7] H. Lee et al., IEEE Photon. Technol. Lett., vol. 31, no. 11, pp. 813-816, 2019.
[8] C. M. Lalau-Keraly et al., Opt. Express, vol. 21, no. 18, pp. 21693-21701, 2013.
[9] T. W. Hughes et al., ACS Photonics, vol. 5, no. 12, pp. 4781-4787, 2018.
[10] A. Dosovitskiy et al., arXiv:2010.11929, 2020.
[11] K. Han et al., IEEE Trans. Pattern Anal. Mach. Intell., vol. 45, no. 1, pp. 87-110, 2023.