MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

Learning-Based THz Multi-Layer Imaging With Model-Based Masks

Wang, Pu; Koike-Akino, Toshiaki; Boufounos, Petros T.; Tsujita, Wataru; Yamashita, Genki

TR2024-017 March 12, 2024

Abstract

This paper demonstrates a learning-based THz multi-layer pixel identification for non-destructive inspection. Specifically, we introduce a recurrent neural network that se- quentially learns features from THz spectrogram segments with masks from model-based sparse deconvolution. Initial performance evaluation on a three-layer sample with contents on all surfaces confirms the effectiveness of the proposed method.

International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz) 2023

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Learning-Based THz Multi-Layer Imaging With Model-Based Masks

P. Wang*, T. Koike-Akino*, P. Boufounos*, W. Tsujita*, G. Yamashita†, T. Fukuta†, and M. Nakajima‡
*Mitsubishi Electric Research Laboratories, Cambridge, MA 02139, USA.
†Mitsubishi Electric Corporation Advanced Technology R&D Center, Amagasaki City, 661-8661, Japan.
‡Institute of Laser Engineering, Osaka University, Osaka 565−0871, Japan.

***Abstract*—This paper demonstrates a learning-based THz multi-layer pixel identification for non-destructive inspection. Specifically, we introduce a recurrent neural network that sequentially learns features from THz spectrogram segments with masks from model-based sparse deconvolution. Initial performance evaluation on a three-layer sample with contents on all surfaces confirms the effectiveness of the proposed method.**

I. Introduction

The use of terahertz (THz) wave for multi-layer material inspection has been demonstrated in [1]–[6] for contactless sensing, operations under adversarial conditions (e.g., fire and smoke), and robustness to dust and dirt.

Nevertheless, inspection results may vary subject to humidity, pixel-to-pixel depth variation due to vibration, and the lack of layer identification. In this paper, we consider THz multi-layer content extraction in non-destructive inspection, screening devices, and quality monitoring systems by utilizing THz penetration capability through non-conducting materials. Specifically, we propose a deep learning-based approach to deal with challenges from 1) depth variation due to the platform vibration and motion; see the raster scanning mode in Fig. 1 (a), 2) shadow effect caused by non-uniform penetrating illumination from front layers to deep layers; see Fig. 1 (b) for example, and 3) the impact of humidity conditions.

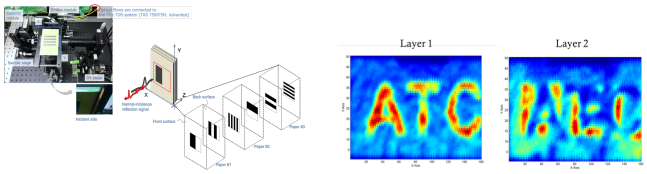
II. Proposed Network Architecture

In Fig. 2, we consider a hybrid approach that leverages data-driven feature learning and model-based deconvolution results.

A. Layer Identification by Sparse Deconvolution

The time-domain THz-TDS waveform can be expressed as a convolution between a sparse depth-wise layer profile $f(t)$ and the reference signal $h(t)$ as $y(t) = h(t) \otimes f(t) = \int_{-\infty}^{\infty} h(\tau) f(t-\tau) d\tau$. By sampling $y(t)$ with a sampling interval T_s, the discrete-time waveform representation is given as $y_n = y(nT_s) = \sum_{m=0}^{M} h_m f_{n-m} + e_n$, where $h_m = h(mT_s)$ and $f_n = f(nT_s)$ are the reference sample and depth profile at corresponding time instances, and e_n is the measurement noise. Equivalently, we can express the sampled waveform in a matrix-vector form as $\mathbf{y} = \mathbf{H}\mathbf{f} + \mathbf{e}$, where $\mathbf{H}$ is the convolution matrix whose rows are cycle-shifted, reversed versions of the reference signal $\mathbf{h}^T$. We employ the following ℓ_1 regularized least square - LASSO - to identify sparse depth profile [7]

$$\hat{\mathbf{f}} = \arg\min_{\mathbf{f}} \frac{1}{2} \|\mathbf{H}\mathbf{f} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{f}\|_1, \tag{1}$$



(a) (b)

Fig. 1. (a) THz-TDS multi-layer imaging with a raster scanning and (b) the shadow of three letters on the 1st layer is clearly shown on the 2nd layer.

where we resort to the fast iterative shrinkage-thresholding algorithm (FISTA) [8]; see Fig. 2 for an illustration.

B. Spectrogram Patch Embedding

To extract time-dependent features, we first represent the time-domain THz-TDS into the time-frequency spectrogram via the short-time Fourier transform (STFT) $Y(t,\omega) = \int y(\tau) g^*(\tau - t) e^{-i\omega\tau} d\tau$, where $g(t)$ is a time-domain localized window function and ω is the frequency variable. Then we divide the spectrogram $|Y(t,\omega)|^2$ into time-dependent 2D patches $\mathbf{P}(t_n) = \{|Y(t,\omega)|^2 | t \in \mathcal{T}(t_n), \omega \in \mathcal{F}(\omega_n)\}$. In our case, we use $\mathcal{T}(t_n) = \{t | t_n - 0.5T_w \le t < t_n + 0.5T_w\}$ where T_w is the window size used in (2) and $\omega_n \in [0, 0.5\omega_s]$ with ω_s is the sampling frequency, rendering a sequence of 2D spectrogram patches sliding over the time (depth) domain $\mathbf{P}_n \triangleq \mathbf{P}(t_n, \omega), n = 1, 2, \cdots, N$. We pass $\mathbf{P}_n$ through a feature extraction $\mathcal{E}(\mathbf{P}_n, \boldsymbol{\theta})$ (e.g., a convolution neural network) parameterized by $\boldsymbol{\theta}$ to produce a latent representation $\mathbf{z}_n \in \mathbb{R}^d$

$$\mathbf{z}_n = \mathcal{E}(\mathbf{P}_n; \boldsymbol{\theta}), \quad n = 1, \cdots, N. \tag{2}$$

C. Recurrent State Update with Model-Based Masks

A standard recurrent neural network, e.g., LSTM, can be used to propagate the above spectrogram patch features over time [9]. Particularly, the RNN is trained to sequentially update time-dependent latent (hidden) variables $\mathbf{h}_n$ using the previous latent variable $\mathbf{h}_{n-1}$ and the current spectrogram feature $\mathbf{z}_n$: $\mathbf{h}_n = \mathcal{R}(\mathbf{h}_{n-1}, \mathbf{z}_n; \boldsymbol{\phi}), n = 1, \cdots, N$, where $\mathcal{R}$ represents an LSTM unit with trainable parameters $\boldsymbol{\phi}$ shared over all time steps. To account for the depth-wise layer structure, we propose to modify the above RNN update step by utilizing the sparse deconvolution depth profile $\hat{\mathbf{f}}$ in (1)

$$\mathbf{h}_n = \mathcal{R}(\mathbf{h}_{n-1}, m_n \mathbf{z}_n; \boldsymbol{\phi}), \quad n = 1, 2, \cdots, N, \tag{3}$$

where $m_n \in \{0, 1\}$ is a scalar mask given by $m_n = \mathbb{1}(|\hat{f}_n| \ge \epsilon)$ with ϵ given as a pre-determined threshold.

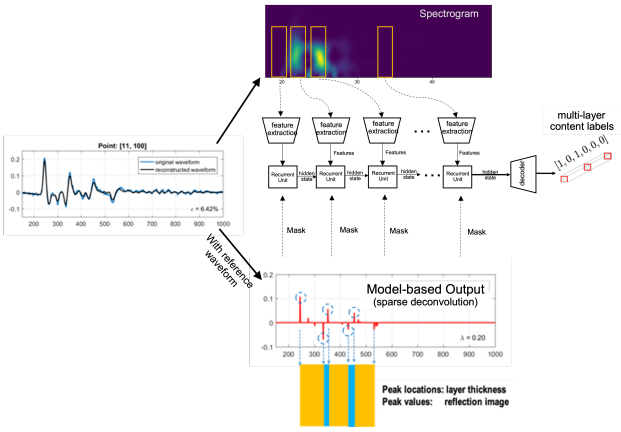


Fig. 2. A THz-TDS multi-layer identification framework that sequentially learns features from THz spectrogram patches and leverages model-based deconvolution results for temporally masking latent features.

D. Decoder

For the decoder, we enforce the last latent variable $\mathbf{h}_N$ to predict the multi-layer binary content labels with a standard multilayer perceptron (MLP) network as

$$\mathbf{u}_n = \mathcal{M}(\mathbf{h}_N; \boldsymbol{\psi}), \tag{4}$$

where $\boldsymbol{\psi}$ consists of the MLP weight matrices and bias terms. We use the (weighted) multi-label binary classification with each label precisely corresponding to a binary label (i.e., $\{0, 1\}$) for each surface. To this end, the output $\mathbf{u}$ is converted to a score vector $\mathbf{s} \in [0, 1]$ using the sigmoid function $s_n = (1 + e^{-u_n})^{-1}$. Then, the total loss takes the weighted average of the N individual losses as

$$L = -\frac{1}{N} \sum_{n=1}^{N} \omega_n c_n \log(s_n), \tag{5}$$

where ω_n is the weight on the n-th surface. The binary imaging result is obtained by comparing $\mathbf{s}$ with a threshold of 0.5.

III. Experimental Results

Fig. 1 (a) shows a three-layer sample mounted on the raster scanning stage of a THz-TDS testbed. Both (front and back) surfaces of each layer are drawn with pencils to cover an area of 40×40 mm^2 that is divided into $8 \times 8 = 64$ patches. Each pixel of the size 5×5 mm^2 corresponds to a unique binary label $\mathbf{c}$. For instance, $\mathbf{c} = [1, 0, 1, 0, 1, 0]$ implies that all front surfaces are covered by the drawing while the back surfaces are blank. With a scanning stepsize of 0.5 mm, we have a set of $10 \times 10 = 100$ THz-TDS waveforms for each pixel and then randomly split them into the training (0.6), validation (0.1), and test (0.3) datasets. The training dataset is augmented by shifting the waveform and adding Gaussian noise to improve the invariance to the depth variation. Fig. 3 shows distinct spectrogram features for THz-TDS waveforms corresponding to different content labels. It is clear that brighter features at the front layers are likely linked to the content covered by the drawing while, at deep layers, it is critical to account for both

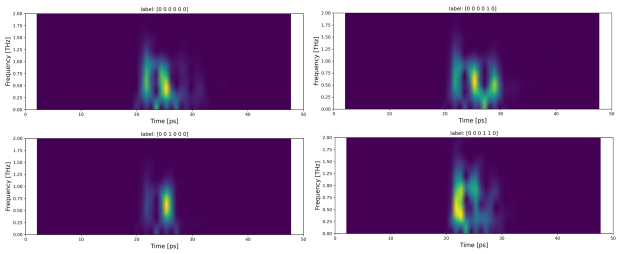


Fig. 3. Distinct spectrogram features for different content labels.

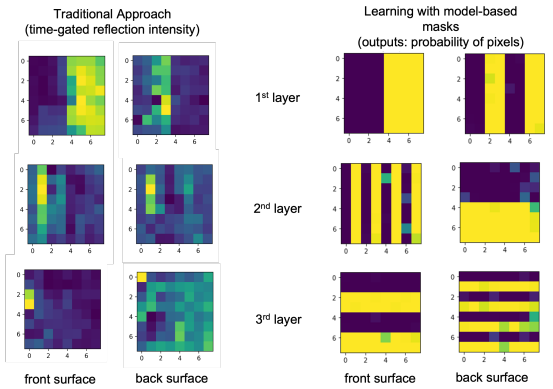


Fig. 4. Comparison between intensity-based and proposed methods.

"current" spectrogram features and propagated latent features from the front layers.

Fig. 4 shows the distorted intensity image of the traditional approach due to aggregated reflection from earlier layers (shadowing effect), the proposed approach provides much cleaner images of score values $\mathbf{s}$ that resemble the true content over all 6 surfaces. The deep layers show slightly higher fluctuations of the score values than the front layers.

IV. Conclusion

This paper proposed a recurrent neural network that sequentially learns features from THz spectrogram segments with masks from model-based sparse deconvolution. Earlier results confirmed its effectiveness.

References

[1] B. Wang, et al., "Metamaterial absorber for THz polarimetric sensing," in *SPIE Photonics West*, 2018.

[2] P. Wang, et al., "Learning-based shadow mitigation for terahertz multi-layer imaging," in *IRMMW-THz*, Sep. 2019.

[3] P. Wang, et al., "Terahertz QR positioning: Experimental results," in *IRMMW-THz*, Nov. 2020.

[4] P. Wang, et al., "Methods and systems for terahertz-based positioning," 2020, US Patent 10,795,151.

[5] P. Wang, et al., "Learning-based THz multi-layer imaging for high-capacity positioning," in *IRMMW-THz*, 2021.

[6] T. Koike-Akino , et al., "Quantum feature extraction for THz multi-layer imaging," in *IRMMW-THz*, 2022.

[7] J. Dong, et al., "Terahertz superresolution stratigraphic characterization of multilayered structures using sparse deconvolution," *IEEE Trans. Terahertz Sci. Technol.*, 2017.

[8] A. Beck and M. Teboulle, "A fast iterative shrinkage thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, 2009.

[9] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, 1997.