Learning-Based Shadow Mitigation for Terahertz Multi-Layer Imaging

P. Wang[†], T. Koike-Akino[†], A. Bose[†§], Rui Ma[†], P. V. Orlik[†],
W. Tsujita[‡], K. Sadamoto[‡] and H. Tsutada[‡], M. Soltanalian[§]
[†]Mitsubishi Electric Research Laboratories, Cambridge, MA 02139, USA.
[‡]Mitsubishi Electric Corporation Advanced Technology R&D Center, Amagasaki City, 661-8661, Japan.
[§]University of Illinois at Chicago, USA

***Abstract*—This paper proposes a learning-based approach to mitigate the shadow effect in the pixel domain for Terahertz Time-Domain Spectroscopy (THz-TDS) multi-layer imaging. Compared with model-based approaches, this learning-based approach requires no prior knowledge of material properties of the sample. Preliminary simulations confirm the effectiveness of the proposed method.**

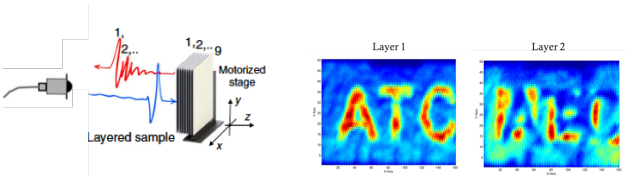
I. Introduction

Over the past years, there has been an increased interest in the use of terahertz wave for gas sensing, non-destructive evaluation, and security screening. In Fig. 1 (a), the time-domain spectroscopy (THz-TDS) system sends an ultra-short pulse (e.g., 1-2 picoseconds) to image layered structures of a sample in a raster scanning mode, thanks to its capability of penetrating a wide range of non-conducting materials.

In the raster scanning mode, the sample under inspection is illuminated by a THz-TDS point source and a programmable mechanical raster moves the sample in the plane perpendicular to the incidental waveform in order to measure the two-dimensional surface of the sample. Single-layer and multi-layer content extractions have been tested in [1]–[4]. One issue here is, due to either irregular sample surfaces or vibration from the mechanical raster move, to deal with depth variation and its induced delay variation from one pixel to another.

Another challenge is to mitigate the shadow effect caused by non-uniform penetrating illumination from front layers to deep layers and consequent non-uniform reflected THz waveforms. This shadow effect has been observed in various THz-TDS imaging results, e.g., Fig. 3 in [2] and Fig. 9 in [5]. Our own experiment on multi-layer hardboard papers, as shown in Fig. 1 (b), also shows clear shadow of the three letters (i.e., 'A', 'T' and 'C') on the first layer onto the three letters (i.e., 'M', 'E' and 'L') of the second layer.

In the literature, only a few efforts were proposed to address this challenge either by assuming *known* wave propagation models from one layer to another [6] or by exploiting the image-domain composition over layers and taking into account the finite alphabet of letters. However, practical factors such as uncertainties on material parameters and device-dependent emitter/receiver specifications significantly impact their performance. The problem of interest here is to mitigate the shadow effect by learning the nonlinear mapping from one layer to another using labelled training data in the pixel domain.



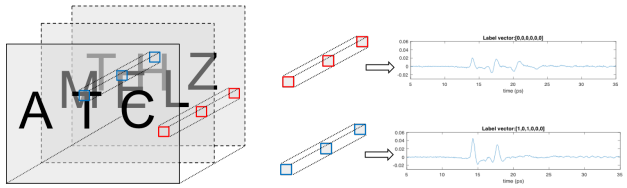
(a) (b)

Fig. 1: (a) THz-TDS multi-layer imaging with a raster scanning (from [2]); (b) our own THz-TDS experiment shows clear shadow of three letters (i.e., 'A', 'T' and 'C') on the 1st layer onto the 2nd layer.

II. Proposed Method

For the THz-TDS imaging, the time-domain waveform is recorded for each pixel at a spot size of the THz emitter and it can be considered as a convolution between the reference waveform and the multi-layer structure. For instance, in Fig. 2 (a), we highlight two pixels (denoted by red and blue squares) over a three-layer sample. Fig. 2 (b) shows reflected time-domain waveform corresponding to the two pixels. For the pixel in red, one can clearly see three positive peaks corresponding to the front surface of each layer and small negative peaks (partially overlapped with positive peaks) corresponding to the back surface of each layer. On the other hand, for the pixel in blue, one can see two dominating peaks due to the stronger reflection on the ink on the first two layers. The third peak almost disappears due to less energy penetrating through the first two layers. Moreover, it is observed that the arrival times of the peaks for the two pixels are slightly different due to depth variation caused by curved sample surface and depth vibration from the mechanical raster scanning.

As a result, for each pixel, we propose to associate the time-domain waveform with corresponding content on the front and



(a) (b)

Fig. 2: THz-TDS time-domain waveforms in the pixel-domain.

978-1-5386-8285-2/19/$31.00 ©2019 IEEE

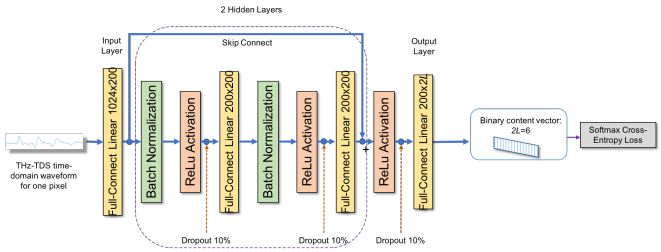


Fig. 3: The proposed DNN architecture for THz-TDS multi-layer imaging.

back surfaces of each layer. Specifically, we use a $2L \times 1$ binary content vector (e.g., $[0,0,0,0,0,0]^T$ for the pixel in red in Fig. 2) to label the L layers, where 1 means that there is a content while 0 denotes no content. By collecting time-domain waveforms for a large number of training pixels, we use the deep neural network (DNN) to learn the nonlinear mapping from the waveform sequence to the content vector.

Fig. 3 shows our proposed DNN architecture for THz-TDS multi-layer imaging. The DNN feeds (truncated) THz-TDS time-domain waveforms and outputs estimated binary content vector. We can also expand the input waveform by including time-domain waveforms from nearby pixels according to the THz-TDS aperture size. Particularly, we truncate the full-length time-domain waveform into a 1024×1 real vector and this waveform vector is fed into the input layer of the DNN, where the input layer first transforms to 200-node dimensions by a fully-connected linear layer. The DNN then employs two hidden layers having 200 nodes per layer, consisting of batch normalization layer, rectified linear unit (ReLU) activation layer with 10% dropout, and fully-connected linear layer. The dropout is a well-established technique to prevent over-fitting for improved generalizability. Additionally, we consider a skip connect jumping from the input of hidden layers to the output of hidden layers in order to learn residual gradient for improved training stability, i.e., ResNet [7]. A fully-connected linear layer following an activation layer with dropout produces the output of the DNN. The DNN is trained to minimize the softmax cross-entropy loss to predict the $2L \times 1$ content vector from labeled training pixels.

III. Synthetic Validation

We verify this learning-based approach using synthetic data generated using the ray-tracing model with random depth variation. Then we apply the learned representation to estimate the content vector of the test dataset corresponds to a sample of three-layer stacked paper. As illustrated in Fig. 2, three letters are plotted on each layer. To generate the reflected time-domain waveform, we use a reference waveform (as shown in Fig. 4 (a) and (b)) from a real-world THz-TDS experiment. The conventional multi-layer imaging is to first identify layer by properly identifying time-domain peaks using peak values and nearby waveforms and then align up the peak magnitude (additional tuning can be done using local spectral features). Fig. 4 (c) shows clear shadow effects on the aligned peak magnitude.

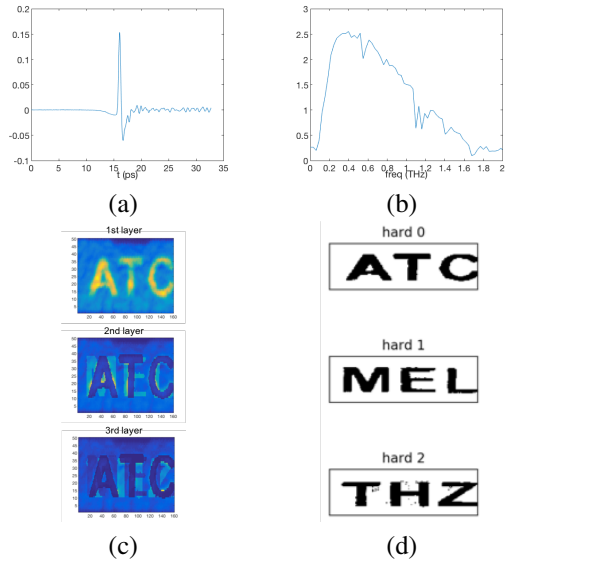


Fig. 4: Synthetic validation using experimental THz reference waveforms ((a) time-domain waveforms and (b) spectrum); (c) Peak magnitude spectrum shows clearly shadow effects and (d) recovered content from the learning-based approach.

On the other hand, for the learning-based approach, the DNN training was performed by adaptive momentum (Adam) stochastic gradient descent method with a learning rate of 0.001, and a mini-batch size of 100. The maximum number of epochs is 500 while early stopping with a patience of 20 was taken place. By placing the estimated content vector in the correct pixel order, the multi-layer imaging results are shown in Fig. 4 (d), where the shadow effect is largely suppressed on the second and third layers. However, one can still notice artifacts on the third layer.

IV. Conclusion

In this paper, we used the proposed DNN architecture to learn the nonlinear mapping from the time-domain waveform to a finite-dimensional binary content vector. Preliminary synthetic validation shows promising results on shadow mitigation and robustness against the depth variation. We plan to include more realistic synthetic modeling (including aperture size, surface granularity and varying sample material properties) and collect real-world THz-TDS data for experimental validation.

References

[1] N. Sunaguchi, et al., "Depth-resolving THz imaging with tomosynthesis," *Optics Express*, vol. 17, no. 12, pp. 9558–9570, June 2009.

[2] A. Redo-Sanchez, et al., "Terahertz time-gated spectral imaging for content extraction through layered structures," *Nature Communications*, vol. 7, pp. 1–7, Sept. 2016.

[3] A. Aghasi, et al., "Sweep distortion removal from terahertz images via blind demodulation," *Optica*, vol. 3, no. 7, pp. 754–762, July 2016.

[4] H. Fu, et al., "Terahertz imaging of binary reflectance with variational Bayesian inference," in *The 43rd IEEE ICASSP*, 2018.

[5] G. C. Walker, et al., "Terahertz deconvolution," *Optics Express*, vol. 20, no. 25, pp. 27230–27241, Dec. 2012.

[6] P. Wang, et al., "Multi-layer Terahertz imaging of non-overlapping contents," in *2018 IEEE SAM*, July 2018.

[7] K. He, et al., "Deep residual learning for image recognition," in *2016 CVPR*, 2016, pp. 770–778.