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

Learning-Based THz Multi-Layer Imaging for High-Capacity Positioning

Wang, Perry; Koike-Akino, Toshiaki; Ma, Rui; Orlik, Philip V.; Yamashita, Genki; Tsujita, Wataru; Nakajima, M.

TR2021-098 September 10, 2021

Abstract

This paper demonstrates a learning-based THz multi-layer pixel identification for contactless three-dimensional (3-D) positioning and encoders. More specifically, we propose a one-dimensional convolution-based residual network to deal with practical issues including 1) depth variation, 2) shadow effect, and 3) content recognition at the back surface of each layer. Experimental validation on a three-layer sample with contents on all surfaces is also provided.

International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz) 2021

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

Learning-Based THz Multi-Layer Imaging for High-Capacity Positioning

P. Wang*, T. Koike-Akino*, Rui Ma*, P. V. Orlik*, G. Yamashita†, W. Tsujita†, 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 considers THz multi-layer imaging for three-dimensional (3D) positioning and encoders. A learning-based approach is introduced to address challenges such as 1) depth variation, 2) shadowing effect, and 3) content recognition at the back surface of each layer. Experimental validation on a three-layer sample with contents on both front and back surfaces is provided to evaluate the effectiveness of the proposed approach.**

I. Introduction

The use of terahertz (THz) wave for absolute positioning has recently been demonstrated in [1]–[8] for contactless sensing, operations under harsh conditions (e.g., fire and smoke), and robustness to dust and dirt. THz barcode positioning using one-dimensional (1D) THz linear encoders was considered and experimentally verified in [1]–[4], while THz QR positioning using two-dimensional (2D) THz encoders was verified in [5]–[8]. In this paper, we extend from THz 1D/2D positioning to THz 3D positioning with layered QR patterns by utilizing THz penetration capability through non-conducting materials.

Several challenges need to be addressed: First, depth variations from one pixel to another due to either the irregular sample surface or the vibration from the mechanical scanning process. Second, the shadowing effect caused by non-uniform penetrating illumination from front layers to deep layers; see Fig. 1 (b). Third, the limited capability to recognize content in the back surface of each layer.

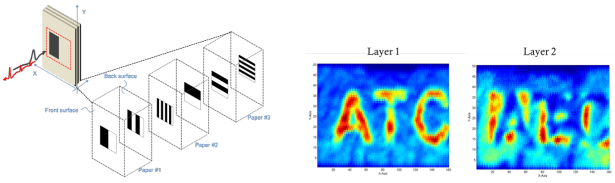
II. Learning-based THz Multi-Layer Positioning

To address the above challenges, we formulate the multi-layer content recovery as a classification problem and introduce a residual learning network to address the depth variation and directly output the score value for each surface.

A. Proposed Network Architecture

The proposed deep neural network architecture for THz multi-layer position is shown in Fig. 2. It first feeds time-domain reflected THz waveform $\mathbf{x}$ to an input block that consists of a convolution layer with a kernel size of 7 and a stride of 2 to reduce the dimension while increasing the number of channels from 1 to 8, a batch normalization (BN) layer for regularization, a rectified linear unit (ReLU) layer for nonlinearity, and a max pooling layer (maxPool):

$$\mathbf{y}_0 = \text{maxPool}\left\{f_0(\mathbf{x}, \boldsymbol{\theta}_0)\right\}, \tag{1}$$



(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 f_0 denotes the input-output mapping that consists of the input convolution layer with kernel parameters $\boldsymbol{\theta}_0$ (weights and biases) and the nonlinear ReLU layer.

Then, $\mathbf{y}_0$ is fed into N_d consecutive residual blocks [9],

$$\mathbf{y}_\ell = f_\ell(\mathbf{y}_{\ell-1}, \boldsymbol{\theta}_\ell) + \mathbf{W}_{\ell-1}\mathbf{y}_{\ell-1}, \quad \ell = 1, 2, \ldots, N_d, \tag{2}$$

where the first term f_ℓ represents the *residual mapping path* with parameters $\boldsymbol{\theta}_\ell$ and the second term denotes the *skip connection path* with weights $\mathbf{W}_{\ell-1}$.

As shown in Fig. 2, all residual blocks have the same structure but with different numbers of channels: The residual mapping path consists of a first 1D convolution layer with a stride of 2 to reduce the input dimension by half, followed by the batch normalization and ReLU layers, and a second convolution layer with the stride of 1; On the other hand, the skip connection path has only one convolution layer (represented by $\mathbf{W}_{\ell-1}$) with a kernel size of 1 and a stride of 2 to match the dimension of the residual mapping path. The outputs from the two paths are added together before feeding into another ReLU activation to generate the output for each residual block. With N_d residual blocks, the network successively downsizes the dimension and increases the number of features.

For the output block, we use an average pooling (avePool) layer to further reduce the dimension, then flatten its output into a vector, and use a fully-connected layer to generate an output vector $\mathbf{u} \in \mathbb{R}^{N\times 1}$ where $N = 2L$ represents the total (front/back) surfaces of L layers. Fig. 2 shows the case of $N = 6$ surfaces of a sample of $L = 3$ layers.

B. Multi-Label Classification

Instead of classifying each input waveform into one of 2^N classes [6], i.e., a multi-class classification, we use the (weighted) multi-label binary classification with each label precisely corresponding to a binary label (i.e., $\{0, 1\}$) for each surface. This multi-label formulation allows us to directly

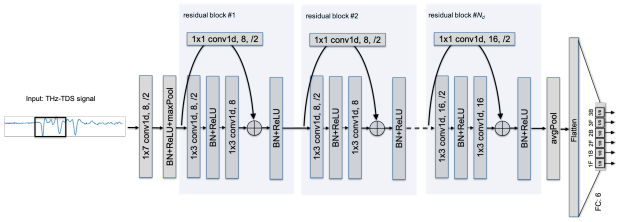


Fig. 2. The THz residual learning neural network with multi-label (multi-surface) binary output classification. For a 3-layer sample, $N = 6$ output scores for all front and back surfaces are computed at the network output.

compute the score for each surface at the output of the neural network. To this end, the output $\mathbf{u}$ is first converted to the score vector $\mathbf{s} \in [0, 1]$ using the sigmoid function

$$s_n = \frac{1}{1 + e^{-u_n}}, n = 1, \cdots, N. \tag{3}$$

For the n-th surface, the binary cross-entropy loss function is computed using the score s_n and the corresponding binary label $c_n \in \{0, 1\}$. Then, the total loss function takes the weighted average of the N individual losses as

$$L = -\frac{1}{N} \sum_{n=1}^{N} \omega_n c_n \log(s_n), \tag{4}$$

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. 3 (a) shows a three-layer sample mounted on the raster scan stage of a THz-TDS testbed. For each layer, both front and back surfaces are drawn by pencils in a way that $8 \times 8 = 64$ pixels in an area of 40×40 mm^2 was scanned. Each pixel of the size 5×5 mm^2 corresponds to a unique $N \times 1$ binary label $\mathbf{c}$. For instance, $\mathbf{c} = [1, 0, 1, 0, 1, 0]$ implies that all front surfaces are covered by the pencil while the back surfaces are blank. With a scanning stepsize of 0.5 mm, we can have a set of $10 \times 10 = 100$ THz-TDS waveforms. We randomly split the collected waveforms into training (0.6), validation (0.1), and test (0.3) datasets. The training dataset is augmented by shifting the waveform (Fig. 3 (b)) and adding Gaussian noise to improve the invariance to the depth variation. The learning trajectory in terms of the total classification accuracy over 6 surfaces is shown in Fig. 3 (c) over epochs. Figs. 3 (d) shows the computed score vectors for a selected pixel in the test dataset which are close to the true multi-label binary label.

To show the final imaging result, we select one testing waveform for one pixel and pull the score vector and final binary results together according to the pixel coordinate. Fig. 4 shows the comparison between the traditional time-gated reflection intensity approach (the left column) and the learning-based approach (the middle and right columns). The traditional approach suffers from the shadowing effect from the front layers to the deep layers and the limited separation between two closely spaced surfaces. On the contrary, the score values of the learning-base method resemble the true content over all 6 surfaces with a reduced shadowing effect. The deep layers show slightly higher fluctuations of the score

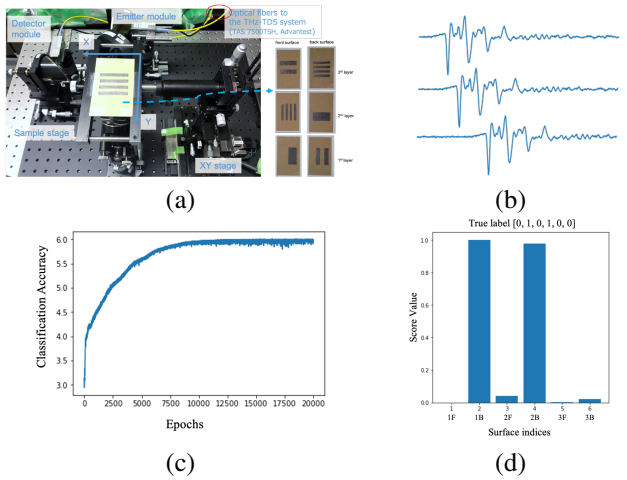


Fig. 3. (a) The THz-TDS testbed with a raster scanning stage and a three-layer sample; (b) examples of augmented training waveforms; (c) training trajectory of classification accuracy over epochs; (d) score values for a selected pixel in the test dataset.

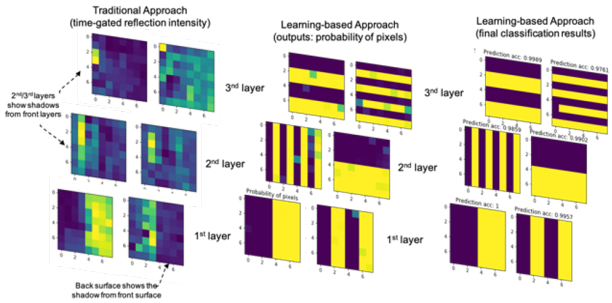


Fig. 4. Performance comparison between the learning-based method and a reflection intensity method.

values than the front layers. By thresholding the score values, the final binary result is shown in the right column.

IV. Conclusion

This paper has shown the learning-based THz 3D content extraction for high-capacity positioning. The residual learning network was experimentally verified to achieve robustness against depth variation, reduce the shadowing effect, and improve the separation between closely spaced surfaces.

References

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

[2] K. Sadamoto, et al., "Terahertz polarimetric sensing for linear encoder," in *IRMMW-THz*, Sep. 2018.

[3] G. Yamashita, et al., "Terahertz polarimetric sensing for linear encoder based on a resonant-tunneling-diode and CFRP polarizing plates," in *IRMMW-THz*, Sep. 2019.

[4] G. Yamashita, et al., "Evaluation of position error of terahertz polarimetric encoder by ray-tracing method," in *IRMMW-THz*, Nov. 2020.

[5] H. Fu, et al., "Terahertz imaging of multi-level pseudo-random reflectance," in *IRMMW-THz*, Sep. 2018.

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

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

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

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