Subsampled terahertz data reconstruction based on spatio-temporal dictionary learning

Vahid Abolghasemi[a], Hao Shen[b], Yaochun Shen[b], Lu Gan[c]

[a]*Faculty of Electrical Engineering, University of Shahrood, Shahrood, 3619995161, IRAN*
[b]*Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, L69 3GL, UK*
[c]*School of Engineering and Design, Brunel University, Uxbridge, UB8 3PH, UK*

Abstract

In this paper, the problem of terahertz pulsed imaging and reconstruction is addressed. It is assumed that an incomplete (subsampled) three dimensional THz data set has been acquired and the aim is to recover all missing samples. A sparsity-inducing approach is proposed for this purpose. First, a simple interpolation is applied to incomplete noisy data. Then, we propose a spatio-temporal dictionary learning method to obtain an appropriate sparse representation of data based on a joint sparse recovery algorithm. Then, using the sparse coefficients and the learned dictionary, the 3D data is effectively denoised by minimizing a simple cost function. We consider two types of terahertz data to evaluate the performance of the proposed approach; THz data acquired for a model sample with clear layered structures (e.g., a T-shape plastic sheet buried in a polythene pellet), and pharmaceutical tablet data (with low spatial resolution). The achieved signal-to-noise-ratio for reconstruction of T-shape data, from only 5% observation was 19 dB. Moreover, the accuracies of obtained thickness and depth measurements for pharmaceutical tablet data after reconstruction from 10% observation were 98.8%, and 99.9%, respectively. These results, along with chemical mapping analysis, presented at the end of this paper, confirm the accuracy of the proposed method.

*Corresponding author
Email address:* `vabolghasemi@ieee.org` (Vahid Abolghasemi)

1. Introduction

Terahertz pulsed imaging (TPI) systems hold great potential in many applications such as medical diagnosis of human tissue, the detection and chemical mapping of illicit drugs and explosives, and pharmaceutical tablet inspection [1, 2, 3]. In a typical TPI measurement, the THz waveform for each pixel is recorded as a function of optical time delay. This provides a three-dimensional (3D) data set where two axes describe an object's spatial content, and the third one represents the (time) depth information. In TPI, the transient electric field, rather than the radiation intensity, is measured. Thus, by applying the one dimensional Fourier transform along the time axis we can get the magnitude and phase information of THz spectral data for each pixel. Despite these advantages, most existing TPI systems suffer from slow scanning speed and high implementation cost due to pixel-by-pixel raster scan mechanism. One approach to speed up the acquisition process is to reduce the number of samples (measurements) and yet preserve the reconstruction quality. Recently, compressive sampling (CS) [4, 5] has been widely used for this purpose. This theory essentially works based on random projections of input signals and sparse representation [6] during reconstruction. It holds great potential for sampling rates reduction, imaging time, power consumption and computational complexity. However, the main challenge in utilizing CS is implementation costs and specifically designing the sampling operator which requires costly and sophisticated hardware modules. Most recent researches on CS-THz has been focused to address this issue. For instance, in [7] the authors propose a spining disk as a plausible sampling operator for high-speed compressive acquisition. In [8], a block-based CS method is proposed, and in [9] a single pixel camera based on Bernoulli random matrix is presented. Recently, an improved version of [8], called adaptive CS, is reported [10]. In this method, additional measurement points are adaptively added at the

2

regions prone to degradation with the aim of improving reconstruction quality.

In this paper, we consider a very simple acquisition scenario, i.e. incomplete (subsampled) data, which can be easily realized without too much of hardware burden. Instead, we aim to utilize advanced reconstruction techniques to retrieve the original data samples with least possible quality degradation. In the sequel, more details are explained.

It should be noted that the spatial and temporal (or spectal) features in 3D TPI data have intrinsic geometrical structures. If these structures are modelled properly, it may lead to new TPI imaging systems with faster acquisition speed and more accurate data analysis. This aspect is an open issue and has not been specifically studied in the THz literature. Over the past few years, there have been increased interests in the study of sparse representation, in which a signal is characterized by a few non-zero coefficients in a certain transform domain. Up until now, sparse modelling has found applications in many image processing-related tasks such as acquistion, denoising, inpainting and super-resolution [11, 12, 13, 14, 15]. Most of existing works are limited to a deterministic sparsifying transform, such as the Fourier, the wavelet and the curvelet etc. More recently, it has been shown that a signal can be represented with fewer coefficients over a learned dictionary from a number of training samples (see [16] and references therein).

In this paper, we aim to investigate dictionary learning from incomplete and noisy 3D TPI data set. In particular, the missing data are first estimated from random subsets using Bicubic interpolation. Incomplete TPI data set can occur in two different scenarios, in which both require retrieving the missing samples; i) result of data transfer from a sender to receiver, and ii) subsampled THz data: a simplified case of CS where merely some samples are missing (this obviously requires much simpler sampling operator than original CS). As a novel tool to address the sample recovery problem, 3D dictionary learning is used here. Spatial-temporal dictionary learning is applied to the 3D data set by exploiting the *joint* sparse model. Finally, denoising is performed based on the learned dictionary through convex optimization. Experimental results show that even

3

with 5% to 20% subsampled data, one can still get reliable spatial, structural and spectral information of the object. To the best of our knowledge, this is the first work demonstrating the advantages of 3D dictionary learning for THz data. These results can be exploited to speed up TPI measurement process substantially by reducing the quantity of data.

This paper is organized as follows. In the next section we describe the proposed method including mathematical expression of the model, dictionary learning and joint sparse recovery. Section 3 is devoted to demonstrating its applicability to the experimental results. Finally, the conclusion is drawn in section 4.

2. Proposed method

2.1. Problem formulation

The first step for recovering the original data from incomplete (subsampled) observations is inpainting. This can be simply achieved using a preliminary and fast approach such as *Bicubic* interpolation [8]. The main drawback is that the results of such techniques are of low quality especially in noisy scenarios. Other inpainting methods which rely on fixed transforms (e.g. 3D wavelet) [17] also do not lead to acceptable results as we will see later in our experimental results. A recent family of methods which try to build up an adaptive dictionary directly over the incomplete data has shown improved results for 2D natural images when the observation rate is above 25% [18]. Efficient extension of these techniques for THz data, especially in severely subsampled noisy data ($< 25\%$) is our objective. To take advantages of existing techniques for both incomplete and noisy data we propose a new model shown in Fig. 1. If we denote the original clean THz datacube by $\mathbf{X} \in \mathbb{R}^{N_x \times N_y \times B}$ and the incomplete noisy datacube by $\underline{\mathbf{Y}}$ (of the same size), the estimated low-quality noisy datacube can be represented by $\mathbf{Y}$ (of the same size). Our aim is to denoise $\mathbf{Y}$ using a dictionary which is specifically designed to exploit both spatial and temporal correlations existing within THz data.
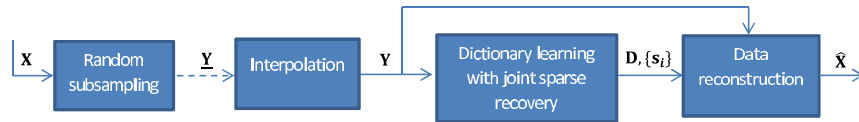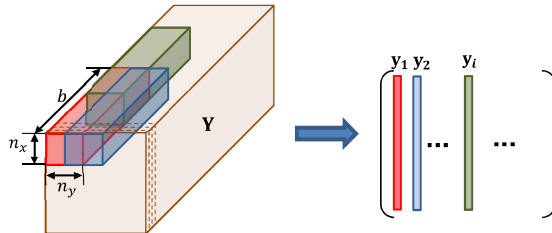
4

Figure 1: Block diagram of the proposed approach.

Figure 2: Geometric illustration of block extraction procedure from the THz datacube.

2.2. Dictionary learning

Since the THz data is huge, one cannot learn a dictionary directly over the
entire datacube. More importantly, 3D nature of THz data should be taken
into account when constituting the training samples. Assume that $\mathbf{y}_i$ is the
i-th training vector which is obtained through partitioning $\mathbf{Y}$ into N small 3D
blocks of size $n_x \times n_y \times b$, where, (n_x, n_y) and b are the block sizes in spatial and
temporal dimensions, respectively. These blocks may have spatial and temporal
overlaps and are always converted to the column-vector $\mathbf{y}_i$ of length $r = n_x n_y b$.
Fig. 2 illustrates the geometrical view of the proposed block extraction. In
addition to capturing spatio-temporal structure in the proposed scheme, we have
the flexibility to define the contributions of spatial and temporal dimensions by
adjusting (n_x, n_y) and b, respectively.

The next step is to find the dictionary $\mathbf{D} \in \mathbb{R}^{r \times k}$ and sparse coefficients,
denoted by $\{\mathbf{s}_i\}_{i=1}^{N}$, from the training set $\{\mathbf{y}_i\}_{i=1}^{N}$. In spite of most traditional
dictionary learning methods in which all training vectors are treated indepen-
dently, we devote a different approach. In order to effectively exploit the 3D
spatio-temporal structure of THz data we cumulate every l training vectors into

5

a matrix $\boldsymbol{\Omega}_j \in \mathbb{R}^{r \times l}$ and solve the following problem:

$$\boldsymbol{\Omega}_j = \mathbf{D}\mathbf{S}_j + \mathbf{E}_j, \tag{1}$$

where j denotes the j-th subset, $\mathbf{D} \in \mathbb{R}^{r \times k}$ is the dictionary, $\mathbf{S}_j \in \mathbb{R}^{k \times l}$ is the corresponding sparse coefficients matrix, and $\mathbf{E}_j \in \mathbb{R}^{r \times l}$ represents the decomposition error. This is a joint sparse model which simultaneously deals with the training vectors in $\boldsymbol{\Omega}_j$. There are different ways to choose the training subset $\boldsymbol{\Omega}_j$. For instance, one can cluster the 3D blocks and then group them based on the amount of similarity within the voxels of these blocks. For simplicity, and yet preserving the sparsity structure of neighboring blocks, we use raster-scanning with full spatial and temporal overlap from left-to-right and top-to-bottom as shown in Fig. 2. We have observed encouraging results using this scheme, but the performance of more sophisticated grouping techniques can be further investigated in the future. Here, we group all the blocks into $\lceil N/l \rceil$ subsets of size l:

$$\underbrace{\{\mathbf{y}_1, \dots \mathbf{y}_l\}}_{\boldsymbol{\Omega}_1}, \underbrace{\{\mathbf{y}_{l+1}, \dots \mathbf{y}_{2l}\}}_{\boldsymbol{\Omega}_2}, \dots \underbrace{\{\mathbf{y}_{N-l+1}, \dots \mathbf{y}_N\}}_{\boldsymbol{\Omega}_{\lceil N/l \rceil}},$$

where $\lceil \cdot \rceil$ is the ceiling operator. Correspondingly, $\mathbf{S}_j$'s are defined as $k \times l$ matrices as follows:

$$\underbrace{\{\mathbf{s}_1, \dots \mathbf{s}_l\}}_{\mathbf{S}_1}, \underbrace{\{\mathbf{s}_{l+1}, \dots \mathbf{s}_{2l}\}}_{\mathbf{S}_2}, \dots \underbrace{\{\mathbf{s}_{N-l+1}, \dots \mathbf{s}_N\}}_{\mathbf{S}_{\lceil N/l \rceil}},$$

and the aim is to solve the following minimization problem subject to sparsity of $\mathbf{S}_j$:

$$\forall j \; \min_{\mathbf{D}, \mathbf{S}_j} \|\boldsymbol{\Omega}_j - \mathbf{D}\mathbf{S}_j\|_F^2 \tag{2}$$

Most typical dictionary learning approaches work based on *"alternating minimization"* [16, 19, 20]. These methods mainly contain two major steps, coefficient update and dictionary update, which are alternately executed until reaching local minima for (2).

In order to jointly update the sparse coefficients we choose MMV (multiple measurement vector) sparse coding framework. In this framework, as opposed
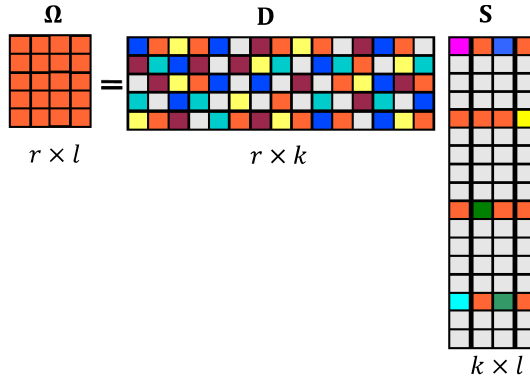
6

Figure 3: MMV model

to SMV (single measurement vector), several sparse vectors are simultaneously recovered. The main assumption in MMV is that the sparse vectors admit a common sparsity pattern, i.e., the locations of non-zeros are the same for all vectors. Fig. 3 is a diagram illustrating the MMV model. There are several algorithms in the literature to solve (2) with respect to $\mathbf{S}_j$. Most of these methods are extended from SMV-based approaches [21, 22, 23]. We have found that SOMP[1] (Simultaneous Orthogonal Matching Pursuit) [22, 23] works well for reconstruction of THz data and hence we utilized it for coefficients update.

Next, for the dictionary update (solving (2) with respect to $\mathbf{D}$), we use a well-established method called K-SVD [16]. K-SVD[2] is a generalization of K-means clustering and updates the dictionary in a column-by-column scheme by using singular value decomposition (SVD) (full details in [16]).

2.3. THz data reconstruction

After finding $\mathbf{D}$ and $\{\mathbf{s}_i\}_{i=1}^N$, we need to estimate the THz data. To do this, we define the following minimization problem:

$$\min_{\mathbf{x}} \lambda \|\mathbf{y} - \mathbf{x}\|_2^2 + \sum_{i=1}^N \|\mathbf{R}_i\mathbf{x} - \mathbf{D}\mathbf{s}_i\|_2^2 + \beta \sum_{i=1}^N \|\mathbf{R}_i\mathbf{x} - \mathbf{m}_i\|_2^2. \qquad (3)$$

---

[1] Available in SPAM: http://spams-devel.gforge.inria.fr/doc/html/doc_spams002.html.
[2] KSVD-Box: http://www.cs.technion.ac.il/~ronrubin/software.html

7

where $\mathbf{y} = \text{vec}(\mathbf{Y})$ and $\mathbf{x} = \text{vec}(\mathbf{X})$, both with length $p = N_x N_y B$, are vectorized representations of datacubes $\mathbf{Y}$ (noisy) and $\mathbf{X}$ (clean), respectively.[3] Also, $\mathbf{y}_i = \mathbf{R}_i \mathbf{y}$ and $\mathbf{x}_i = \mathbf{R}_i \mathbf{x}$, where $\mathbf{R}_i$ is a huge binary matrix of size $r \times p$. $\mathbf{R}_i$ has only one non-zero (i.e. 1) in each row. As a block extraction operator, $\mathbf{R}_i$ extracts the voxels belonging to i-th block.

In (3), the leftmost term is the error between noisy and clean data, the middle term is related to the sparse decompostion error, and the rightmost term is the smoothness constraint with regularization parameter β. Moreover, $\mathbf{m}_i = [m_i, m_i, \ldots]^T$ is the $r \times 1$ mean vector of the i-th block (i.e. average of all elements in $\mathbf{R}_i \mathbf{x}$). Problem (3) is convex in $\mathbf{x}$ and can be solved by zeroing its gradient, with respect to $\mathbf{x}$, which finally leads to:

$$\widehat{\mathbf{x}} = \left(\lambda \mathbf{I} + (1 + \beta) \sum_{i=1}^{N} \mathbf{R}_i^T \mathbf{R}_i \right)^{-1} \left(\lambda \mathbf{y} + \sum_{i=1}^{N} \mathbf{R}_i^T (\mathbf{D}\mathbf{s}_i + \beta \mathbf{m}_i) \right). \qquad (4)$$

Although the above expression seems computationally expensive at the first glance, it does not need to be directly calculated in practice. Instead, since the inverting term in (4) is diagonal, we obtain the estimated datacube in a voxel-wise fashion (similar to the strategy used in [24]). We can show the voxel-wise calculation of one voxel $\hat{x}$ (taking the corresponding voxels in $\mathbf{y}$ and $\mathbf{m}$ into account) using the following sets of equations:

$$\min_{x} \left\{ \mathcal{C}(x) = \lambda(x - y)^2 + (x - ds)^2 + \beta(x - m)^2 \right\} \qquad (5)$$
$$\rightarrow \frac{\partial \mathcal{C}}{\partial x} = 2\lambda(x - y) + 2(x - ds) + 2\beta(x - m) = 0$$
$$\rightarrow \hat{x} = \frac{1}{1 + \beta + \lambda} (\lambda y + ds + \beta m)$$

where $\mathcal{C}$ is the cost function to be minimized, and ds is the corresponding voxel in $\mathbf{Ds}$. It is seen that the last equation above well matches with (4) which is the non-practical form of reconstructing datacube.

The pseudo-code of the proposed method is given in Algorithm 1.

---

[3]Note that $\mathbf{x}$ and $\mathbf{y}$ should not be confused with 3D blocks where always have subscript indices thorough the paper.

Algorithm 1: Pseudo-code of the proposed method.

Input: Incomplete noisy datacube $\underline{\mathbf{Y}}$.

Output: Recovered datacube $\widehat{\mathbf{X}}$.

1 begin

2   *Interpolation*:

3   Apply *Bicubic* interpolation on $\underline{\mathbf{Y}}$ and yield $\mathbf{Y}$;

4   *Dictionary learning*:

5   Extract the spatio-temporal blocks from $\mathbf{Y}$ and obtain $\{\mathbf{y}_i\}_{i=1}^N$;

6   Group the blocks and constitute $\{\mathbf{\Omega}_j\}_{j=1}^{\lceil N/l \rceil}$;

7   Initialize the dictionary $\mathbf{D}$ with the DCT (discrete cosine transform) basis;

8   **repeat**

9   Apply SOMP to $\mathbf{\Omega}_j$ and yield $\mathbf{S}_j$ for all $j = 1, \ldots \lceil N/l \rceil$;

10   Update $\mathbf{D}$ using K-SVD dictionary update [16];

11   **until** *stopping criterion is met*;

12   *THz data reconstruction*:

13   Calculate (4) to yield $\widehat{\mathbf{x}}$ and then reshape it to $\widehat{\mathbf{X}}$;

14 end

---

3. Experimental results

We present the results of applying the proposed method to two different types of THz data. The first dataset, which we call it *"T-shape"*, has size $200 \times 200 \times 512$, and acquired across an area of 20mm×22mm using a TPIscan-1000 system (TeraView Ltd, Cambridge, U.K.), covering a spectral range from 0.1 to 3.5 THz. The sample used is a polythene pellet of a diameter of 25 mm. Inside the pellet there is a T-shaped plastic sheet which locates approximately 0.2 mm below the sample surface. For this dataset, the *structural* information such as thickness and depth are of interest. The second data is a low spatial resolution data called *"Tablet data"* acquired from pharmaceutical tablets. Two of such sets (called LA and TP) have size $42 \times 42 \times 512$, and other seven sets have size $49 \times 49 \times 609$. Extraction of *spectral* information such as chemical mapping is crucial for this dataset.

(a)   (b)

SNR = 16.21 dB   SNR=16.23 dB   SNR = 19.02 dB

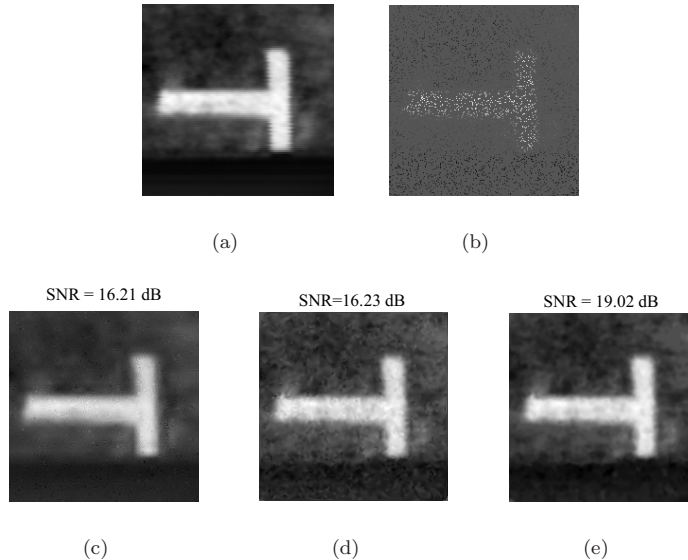(c)   (d)   (e)

Figure 4: T-shape data at one temporal band: (a) original image, (b) 10% incomplete noisy observations. Reconstruction results using (c) soft-thresholding in 3D wavelet domain, (d) *Bicubic* interpolation, and (e) proposed method.

3.1. T-Shape data

Following the proposed model, we first applied *Bicubic* interpolation to the noisy incomplete data, yielding $\mathbf{Y}$. We considered input additive Gaussian noise, leading to input SNR in the range $15 \sim 20$ dB, which is a reasonable range in practice. We then applied the proposed dictionary learning method to find the dictionary and corresponding sparse coefficients to reconstruct the entire datacube. Due to importance of structural information in T-shape data, we gave more contribution to the spatial domain than temporal domain for the dictionary learning. Hence, the experiments were conducted by setting the spatial size $n_x = n_y = 8$, and two different temporal sizes: $b = 1$ (spatial dictionary), and $b = 4$ (spatio-temporal dictionary). In any case we used full spatial and temporal overlaps. According to these settings, the obtained dictionaries were of size 64×256 and 256×256. A sample reconstructed image with $b = 4$, $l = 10$, $\lambda = 0.5$, $\beta = 0.1$, and input SNR of 17 dB is given in Fig. 4. It is seen from
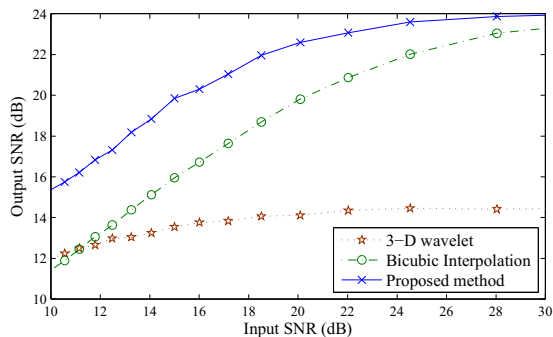
10

Figure 5: T-shape data: SNR of reconstructed data at 10% observation rate versus input SNR.

this figure that the proposed method has recovered and denoised the original image from only 10% of noisy samples with highest SNR among other methods. The proposed method was able to enhance the SNR of *Bicubic* interpolation results by up to 2.5 dB. Other methods, i.e. soft-thresholding after applying 3D *symlet4* wavelet transform [17], shows weaker performances compared to the proposed method. We also show in Fig. 5 the reconstructed SNR at the fixed observation rate of 10% versus different input SNRs. As seen from this figure, the proposed approach outperforms other methods.

Next, performance of the proposed method against different values of b and l is investigated. For this purpose, SNRs of the recovered datacube for different observation rates are given in Table 1. From this table, the advantage of spatio-temporal dictionary can be observed by comparing the SNRs at $b = 4$, with those at $b = 1$. Also, the given SNRs in this table indicates that the proposed method is more effective than 3D wavelet transform. We empirically observed that $l \approx 10$ gives the best performance for T-shape dataset.

Table 2 gives an insight about the computational complexity of the proposed method. In this table, the elapsed time of the dictionary learning stage for different selections of b and l is shown. It is clearly seen that large values of l can significantly reduce the computation time. This behavior supports the advantages of using joint sparse recovery which has already shown to improve

11

Table 1: T-shape data: recovery accuracy (SNR) in dB for different methods. The results are shown for several observation rates, and different selections of b and l.

observation rate	SNR (dB)			b, l
	3D wavelet	*Bicubic* interp.	proposed method	
5%	10.53	16.89	17.45	1, 1
			17.69	1, 10
			18.78	4, 1
			18.96	4, 10
10%	14.71	17.68	19.41	1, 1
			19.46	1, 10
			20.17	4, 1
			20.94	4, 10
15%	16.72	17.87	20.05	1, 1
			20.33	1, 10
			20.86	4, 1
			22.25	4, 10
20%	17.69	17.96	20.88	1, 1
			21.03	1, 10
			22.53	4, 1
			23.15	4, 10

Table 2: T-shape data: The computation time of dictionary learning step in the proposed method for different selections of b and l. These times are in second and calculated per frame.

observation rate / b, l	10%	20%
1,1	110.3	122.5
1, 10	5.0	7.6
4, 1	93.3	97.6
4, 10	4.8	5.9

the quality as well (in Table 1). It can be further seen from Table 2 that increasing b also decrease the computation time, though not as dramatically as that for l.

As a useful measure, we performed the *thickness* and *depth* evaluations before and after reconstruction of T-shape object. These results give structural information about the object of interest. Fig. 6 (a) geometrically illustrates
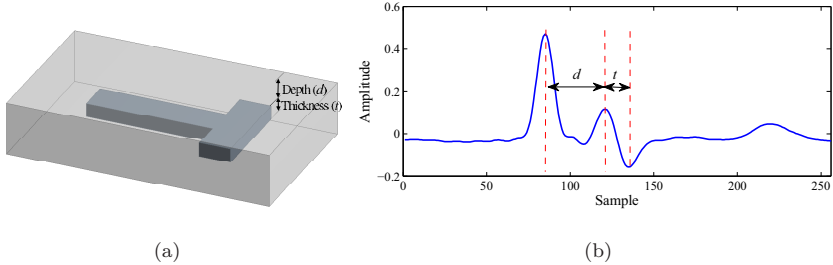
Figure 6: T-shape data: (a) geometrical representation of T-shape object, and (b) sample time-domain waveform at one spatial location.

Table 3: T-shape data: measured "mean±standard deviation" of thickness and depth for different observation rates.

	observation rate	reconstructed	original
Thickness (mm)	5%	0.0766 ± 0.0063	0.0750 ± 0.0029
	10%	0.0741 ± 0.0063	
	15%	0.0756 ± 0.0056	
	20%	0.0748 ± 0.0054	
Depth (mm)	5%	0.1889 ± 0.0079	0.1902 ± 0.0071
	10%	0.1909 ± 0.0087	
	15%	0.1911 ± 0.0079	
	20%	0.1899 ± 0.0082	

these two parameters. Thickness (t) and depth (d) calculations are practically useful to identify inhomogeneities or defects in the object of interest [1]. In order to find these parameters we refer to a sample temporal waveform at one spatial location shown in Fig. 6 (b). The obtained values for these two parameters are subject to appropriate scalings to be converted to millimeter (more details can be found in [1]). The calculated means and standard deviations of these parameters for different observation rates are given in Table 3. The results in this table are shown for both original and reconstructed data using the proposed method. It is found from Table 3 that the accuracy of reconstruction using the proposed method is very high, as the depth and thickness are approximately equivalent to the original data.
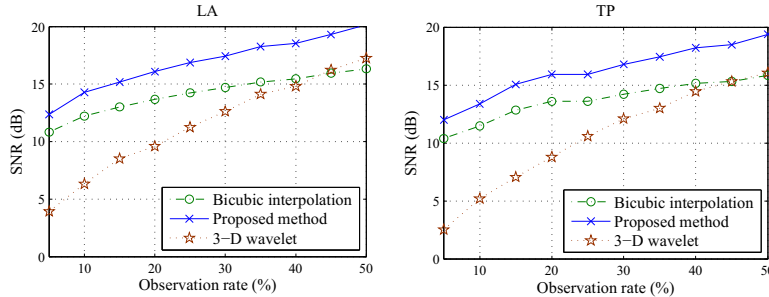
13

Figure 7: Tablet data: the reconstructed SNR versus observation rate for both LA and TP datasets at 20 dB input SNR.

3.2. Tablet data

The dictionary learning settings for Tablet data is different from those of T-Shape data. Due to very low spatial resolution of Tablet data, it is natu- ¹⁹⁵ ral to give a much higher contribution to the temporal dimension during the dictionary learning. This is inline with the fact that the temporal/spectral information is more valuable than spatial information in pharmaceutical studies [1][25]. Therefore, we selected $n_x = n_y = 2$ and $b = 128$ with full spatial and temporal overlaps within the blocks.

²⁰⁰ As a quantitative measure, we show in Fig. 7 the variations of reconstructed SNR of both LA and TP datasets versus different observation rates, where $l = 10$, $k = 64$, $\lambda = 0.5$ and $\beta = 0.2$ were used for the proposed method. These values are chosen empirically and are manually tuned to achieve the best performance among other selections. We also observed that the achieved results ²⁰⁵ are not sensitive to variations of λ and β around the selected values. The dictionary columns size, i.e. k, is set to 64 due to lower spatial resolution in Tablet data compared to T-shape data. The input SNR for this experiment was set to 20 dB. Similarly, we show the reconstructed SNR versus input SNR at 20% observation rate. It is clearly seen that the proposed method is able to ²¹⁰ significantly improve the reconstructed SNR in both figures.

In order to observe the robustness of the proposed approach, more sets of tablet data were included in our experiments and the average performance was
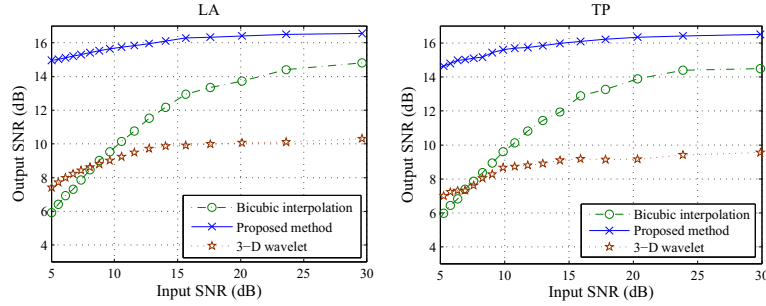
14

Figure 8: Tablet data: the reconstructed SNR versus input SNR for both LA and TP datasets at 20% observation rate.

measured. Seven incomplete Tablet datasets at different observation rates were reconstructed by using well-known methods. Three input SNR levels were used ₂₁₅ in this experiment. The results are given in Table 4. It is seen that the achieved results using the proposed method is better than those obtained using conventional techniques. Moreover, the average output SNRs are compatible with what obtained in the previous experiment (Figures 7 and 8).

As mentioned before, chemical mapping is a useful tool to analyze and observe the uniformity of the tablet [1]. Chemical mapping can be obtained using the terahertz spectral information. This evaluation should be performed in

Table 4: Tablet data: output SNRs (dB) at different observation rates and input SNRs (dB), averaged over seven datasets.

Input SNR (dB)	observation rate (%)	Output SNR (dB)		
		3D wavelet	*Bicubic* interp.	proposed method
	10	7.91	9.49	14.71
10	20	9.31	11.14	16.26
	30	12.16	13.75	18.12
	10	8.34	12.40	16.13
20	20	10.67	14.83	17.22
	30	12.24	15.91	19.49
	10	9.81	14.92	18.23
30	20	12.64	16.78	20.12
	30	14.31	18.63	22.70

15

the frequency-domain, thus, we first take Fourier transform of the time-domain waveforms for both original and reconstructed data. Then, a reference (spectral) waveform denoted by $\mathbf{u}$ is selected from original LA dataset at one pixel location (normally from the middle of the tablet). After that, a spectral matching should be performed to generate the chemical map. We used the cosine correlation mapping (CCM) [26], as used in [1], for this purpose. If we define a spectral frequency-domain vector by $\mathbf{v}$, then the CCM can be calculated via:

$$\cos(\theta) = \sum_{p=1}^{P} \frac{u_p \cdot v_p}{\|\mathbf{u}\|_2 \cdot \|\mathbf{v}\|_2}, \tag{6}$$

where P is the number of spectral bands. Parameter θ is the angle between $\mathbf{v}$ and $\mathbf{u}$, and therefore, $\cos(\theta)$ is between 0 and 1. This evaluation should be performed for all pixels of the corresponding dataset. Smaller angle (larger $\cos(\theta)$) means better match between reference spectra and the waveform under evaluation. The results of chemical mapping for our tablet data is shown in Fig. 9. We calculated CCM for four cases; TP with LA reference, LA with LA reference, TP with TP reference, and TP with LA reference. It is seen from Fig. 9 that both the reconstructed data have a close chemical map similar to the original ones, and that they have a smooth and uniform shape.

4. Discussion and Conclusion

In this paper, we addressed the THz data recovery from an incomplete noisy set of observations. The main advantage of starting with incomplete (subsampled) THz data is fast and inexpensive acquisition process. Inspired by recent developments in dictionary learning and sparse recovery frameworks, we proposed a spatio-temporal dictionary learning technique to find an efficient sparse domain for the THz data. The proposed method aimed at exploiting the existing temporal and spatial correlations. It is worthy here to know the difference between the time-domain THz system and visible-light camera by noting their image formation details. In fact, even for a static object, one can still get 3D THz data, while for a visible-light camera, the 3D data leads to a video sequence
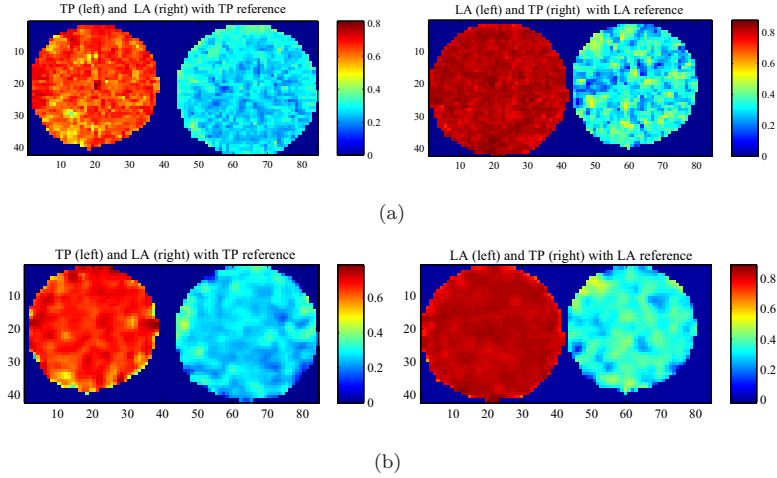
16

Figure 9: Tablet data: the chemical map for (a) original and (b) reconstructed data from 20% observation rate.

with moving scene. To be more accurate, in terms of spatial correlation, THz signal exhibits some similar correlations as that obtained by a visible camera. However, in terms of temporal-correlation, it is significantly different from that of video sequences at visible spectrums. Note that the temporal THz data will either give about structural information like thickness and depth or provide some spectral information, e.g., the chemical component of an object, as presented in the previous section. Extensive sets of experiments were conducted to support the effectiveness and accuracy of the proposed method. The thickness and depth calculations, and chemical mapping analysis for two different types of THz data, along with the achieved SNRs, confirmed the advantages of the proposed approach. In future, we are going to extend the proposed method for the complex scenario where a joint spatio-temporal and frequency domain dictionary learning can be achieved.

5. Acknowledgement

17

References

[1] Y. C. Shen, P. Taday, Development and application of terahertz pulsed imaging for nondestructive inspection of pharmaceutical tablet, IEEE Journal of Selected Topics in Quantum Electronics 14 (2) (2008) 407 –415.

[2] R. K. May, M. J. Evans, S. Zhong, I. Warr, L. F. Gladden, Y. Shen, J. A. Zeitler, Terahertz in-line sensor for direct coating thickness measurement of individual tablets during film coating in real-time, Journal of Pharmaceutical Sciences 100 (4) (2011) 1535–1544.

[3] J. Muller, D. Brock, K. Knop, J. A. Zeitler, P. Kleinebudde, Prediction of dissolution time and coating thickness of sustained release formulations using Raman spectroscopy and terahertz pulsed imaging, European Journal of Pharmaceutics and Biopharmaceutics 80 (3) (2012) 690–697.

[4] D. L. Donoho, Compressed sensing, IEEE Transactions on Information Theory 52 (4) (2006) 1289 –1306.

[5] E. J. Candes, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, IEEE Transactions on Information Theory 52 (2) (2006) 489 – 509.

[6] E. J. Candes, M. B. Wakin, An introduction to compressive sampling, IEEE Signal Processing Magazine 25 (2) (2008) 21 –30.

[7] H. Shen, L. Gan, N. Newman, Y. Dong, C. Li, Y. Huang, Y. C. Shen, Spinning disk for compressive imaging, Optics Letters 37 (1) (2012) 46–48.

[8] C. Sang-Heum, L. Sang-Hun, N.-G. Chan, O. Seoung-Jun, S. Joo-Hiuk, P. Hochong, A. Chang-Beom, Fast terahertz reflection tomography using block-based compressed sensing, Optics Express 19 (17) (2011) 16401–16409.

18

[9] W. L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, D. M. Mittleman, A single-pixel terahertz imaging system based on compressed sensing, Applied Physics Letters 93 (12) (2008) 121105–121105–3.

[10] K. Kim, D.-G. Lee, W.-G. Ham, J. Ku, S.-H. Lee, C.-B. Ahn, J.-H. Son, H. Park, Adaptive compressed sensing for the fast terahertz reflection tomography, IEEE Transactions on Terahertz Science and Technology 3 (4) (2013) 395–401.

[11] Z. Xu, J. Sun, Image inpainting by patch propagation using patch sparsity, IEEE Transactions on Image Processing 19 (5) (2010) 1153–1165.

[12] M. Bertalmio, L. Vese, G. Sapiro, S. Osher, Simultaneous structure and texture image inpainting, IEEE Transactions on Image Processing 12 (8) (2003) 882–889.

[13] T. Ruzic, A. Pizurica, Context-aware patch-based image inpainting using markov random field modeling, IEEE Transactions on Image Processing 24 (1) (2015) 444–456.

[14] M. Elad, A. Feuer, Restoration of a single superresolution image from several blurred, noisy, and undersampled measured images, IEEE Transactions on Image Processing 6 (12) (1997) 1646–1658.

[15] S. Villena, M. Vega, R. Molina, A. Katsaggelos, A non-stationary image prior combination in super-resolution, Digital Signal Processing 32 (0) (2014) 1 – 10.

[16] M. Aharon, M. Elad, A. Bruckstein, K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation, IEEE Transactions on Signal Processing 54 (11) (2006) 4311–4322.

[17] A. Woiselle, J.-L. Starck, J. Fadili, Inpainting with 3D sparse transforms, in: the SPIE Wavelets XIII, Vol. 7446, 2009, pp. 74461C–74461C.

[18] M. Elad, Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing, Springer, 2010.

[19] K. Engan, S. O. Aase, J. Hakon Husoy, Method of optimal directions for frame design, in: IEEE International Conference on Acoustics, Speech, and Signal Processing –ICASSP'99, 1999, pp. 2443–2446.

[20] B. A. Olshausen, D. J. Field, Emergence of simple-cell receptive field properties by learning a sparse code for natural images, Nature 381 (6583) (1996) 607–609.

[21] S. F. Cotter, B. D. Rao, K. Engan, K. Kreutz-Delgado, Sparse solutions to linear inverse problems with multiple measurement vectors, IEEE Transactions on Signal Processing 53 (7) (2005) 2477 – 2488.

[22] J. Tropp, A. Gilbert, M. Strauss, Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit, Signal Processing 86 (3) (2006) 572–588.

[23] J. Tropp, Algorithms for simultaneous sparse approximation. Part II: Convex relaxation, Signal Processing 86 (3) (2006) 589–602.

[24] M. Elad, M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, IEEE Transactions on Image Processing 15 (12) (2006) 3736 –3745.

[25] R. Cogdill, R. Forcht, Y. C. Shen, P. Taday, J. Creekmore, C. Anderson, J. Drennen, Comparison of terahertz pulse imaging and near-infrared spectroscopy for rapid, non-destructive analysis of tablet coating thickness and uniformity, Journal of Pharmaceutical Innovation 2 (2007) 29–36.

[26] J. Schwarz, K. Staenz, Adaptive threshold for spectral matching of hyperspectral data, Canadian Journal of Remote Sensing 27 (2001) 216–224.