

Journal Article (Accepted Manuscript Version)

Coarse-to-Fine Sparse 3D Reconstruction in THz Light Field Imaging

Abdulraouf Kutaish, Miguel Heredia Conde, Ullrich Pfeiffer

This document is the accepted manuscript version that has been published in final form in:

IEEE Sensors Letters

DOI: <https://doi.org/10.1109/LSENS.2024.3454567>

© 2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Persistent identifier of this version: <https://doi.org/10.25926/23t7-gn85>

Sensor Applications

Coarse-to-Fine Sparse 3D Reconstruction in THz Light Field Imaging

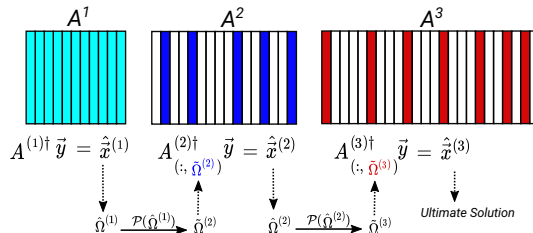
Abdulraouf Kutaish, Miguel Heredia Conde*, and Ullrich Pfeiffer**

Institute for High-Frequency and Communication Technology (IHCT), Wuppertal University
Rainer-Gruenter-Str. 21, 42119 Wuppertal

* Member, IEEE

** Fellow, IEEE

Abstract—THz light field imaging inherently allows capturing the 3D geometry of a target but at the cost of an increased data volume. Compressive sensing techniques are instrumental in minimizing data acquisition requirements. However, they often rely on computationally expensive sparse reconstruction approaches with high memory footprint. This research introduces an advanced coarse-to-fine (CTF) sparse 3D reconstruction strategy aimed at enhancing the precision of reconstructed images while significantly reducing computational load and memory footprint. By employing a sequence of sensing matrices of increasing resolution, our approach avoids falling into an ill-conditioned inversion and strikes a balance between reconstruction quality and computational efficiency. We demonstrate the effectiveness of this CTF strategy through its integration with several established algorithms, including Basis Pursuit (BP), Fast Iterative Shrinkage-Threshold Algorithm (FISTA), and others. The results showcase the potential of the CTF approach to improve 3D image reconstruction accuracy and processing speed in THz light field imaging.



Index Terms—Coarse-to-Fine Approach, THz Light-Field Imaging, Compressive Sensing, Computational Imaging, Sparse reconstruction.

I. INTRODUCTION

Terahertz (THz) imaging has emerged as a pivotal technology in non-invasive, high-resolution diagnostics and analysis, surpassing the capabilities of traditional visible and infrared light imaging [1]. THz radiation, unlike other imaging modalities, can penetrate opaque materials like plastics, ceramics, and clothing, making it uniquely valuable in various applications, from security screening to industrial component inspection [2]. 3D THz imaging is highly effective in the security, biomedical, and industrial sectors [3]. Techniques such as triangulation and backpropagation enable rapid 3D image reconstruction from THz data, though they compromise precision [4]. Addressing the computational challenges caused by the vast data volumes and the large ambient dimension arising from the uniform 3D grid is crucial for improving accuracy [5].

Sparse reconstruction algorithms, key to compressive sensing [6], enhance THz imaging efficiency by reducing data volume through sparsity [7]. However, although they decrease sensing demands, they do not lessen the computational effort needed to solve the inverse problem, highlighting the need for advanced techniques.

This paper introduces a novel form of optimization of sparse reconstruction algorithms through a coarse-to-fine (CTF) approach, building on the foundations laid in [8], [9]. The proposed CTF scheme is illustrated in the graphical abstract. It starts with a coarse reconstruction to identify regions of interest, leveraging a low-resolution sensing matrix, \mathbf{A}^1 , to solve the optimization problem. The initial estimated solution $\hat{\vec{x}}^{(1)}$ is refined using Otsu's thresholding [10] to define a support set $\hat{\Omega}^{(1)}$. This support is propagated to the next scale using the mapping \mathcal{P} , resulting in the support for the next step $\hat{\Omega}^{(2)}$. This process is repeated until the finest scale is reached, corresponding to the full-resolution sensing matrix. We applied this strategy across various algorithms, including Basis Pursuit (BP), Basis

Pursuit Denoising (BPDN), and others, demonstrating its effectiveness in enhancing the precision and efficiency of sparse reconstructions.

THz light-field data, which captures both the intensity and direction of THz rays allowing for dynamic refocusing and perspective changes, is the ongoing focus of our research group. However, in this paper, we consider solely the 3D-resolved intensity aspect of the light-field data, treating it in the same way as traditional volumetric data. Consequently, the CTF approach developed for THz light-field data can also be applied to other forms of volumetric data, such as MRI and CT scans [11], by respective modification of the sensing model.

II. RELATED WORK

Recent advances in sparse reconstruction techniques have significantly impacted high-resolution imaging. A noteworthy development in this field is the Coarse-to-Fine (CTF) estimation method proposed in [12], which enhances the capabilities of inverse synthetic aperture radar (ISAR) imaging. Building upon the principles of CS, this method optimizes the imaging process by progressively discarding columns from the sensing matrix as the reconstruction transitions from coarser to finer scales.

The novel strategy of substituting a single large sensing matrix with a sequence of smaller ones significantly reduces the computational load. Building upon this concept, we explore similar optimization strategies in THz light-field imaging, widening the range of methods considered to solve the resulting sub-problems, from conventional least squares to sparsity-promoting approaches.

A similar approach had previously been suggested in [9] for jointly reconstructing sets of Time-of-Flight (ToF) raw images, leveraging the multi-scale nature of wavelets within a rank-aware Order Recursive Matching Pursuit (ORMP) [13] sparse reconstruction procedure. The *tree structure* of 1D signals in multiscale bases such as wavelets had initially been exploited to improve the performance of Orthogonal Matching Pursuit (OMP) in [8].

Corresponding author: Abdulraouf Kutaish (e-mail: kutaish@uni-wuppertal.de).

III. METHODOLOGY AND FOUNDATIONS

A. THz System Architecture and Design

As shown in Fig. 1, our THz imaging system features a 3×3 matrix of independently operating TicWave TWS-ID02 THz cameras. Each camera offers up to 50 FPS, a 0.3-1.1 THz bandwidth, 1G counts/W responsivity, and 10-20 pW/ $\sqrt{\text{Hz}}$ sensitivity at 320 GHz. With a 33×33 pixel Focal Plane Array (FPA) per aperture, the system captures up to $m = 9 \times 33 \times 33$ measurements in a single shot. We consolidate these measurements from all apertures into a single column vector, $\vec{y} \in \mathbb{R}^m$. For more details, see section IV.

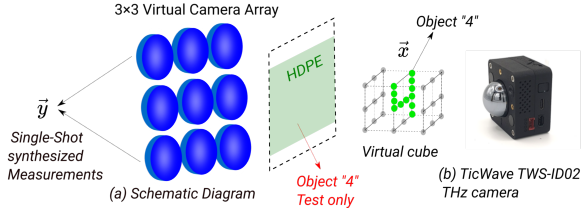


Fig. 1: (a) Schematic overview of the multi-aperture-based system operation; (b) The primary camera unit utilized in the system.

B. Sensing Matrix Configuration

In our THz imaging system, the measurement vector $\vec{y} \in \mathbb{R}^m$ represents the direct output from the sensing process. The original scene is visualized as a 3D cubic volume uniformly discretized with dimensions $R_x \times R_y \times R_z$. The discretization step size is $\Delta x = \Delta y = \Delta z$, which results in a 3D tensor $X \in \mathbb{R}^{n_x \times n_y \times n_z}$, where $n_y := R_y / \Delta y$, $\forall \gamma \in \{x, y, z\}$. This tensor X is then reshaped into the vector $\vec{x} \in \mathbb{R}^n$, where $n := \prod_{\gamma} n_{\gamma}$ represents the total number of discretization points in the volume. The linear sensing model is defined by the equation:

$$\vec{y} = \mathbf{A}\vec{x} + \vec{n}, \quad (1)$$

where $\mathbf{A} \in \mathbb{R}^{m \times n}$ is the sensing matrix that represents a linear transformation of the scene's discretized volume \vec{x} into the data space of the measurements \vec{y} . The vector \vec{n} denotes the noise in the measurements. Each column of \mathbf{A} , denoted \vec{a}_k , represents the response pattern of all pixels to each corresponding element $x_k = \vec{x}[k]$ in the 3D cubic volume. This response is empirically determined by placing a point THz source at x_k and recording the readings from all pixels, yielding a total of $m = 9 \times 33 \times 33$ data points.

C. Computational Imaging Algorithms

The primary goal in THz 3D imaging reconstruction is to extract the true signal \vec{x} from the noisy light field data \vec{y} . This involves solving the inverse problem defined in (1). Several computational algorithms have been employed to address this challenge, each designed to solve mathematical challenges inherent in extracting the true signal \vec{x} from the measurement vector \vec{y} .

The so-called *right* Moore–Penrose pseudoinverse (PI) method [14] is utilized to solve the linear least squares problem:

$$\min_{\vec{x}} \|\mathbf{A}\vec{x} - \vec{y}\|_2^2 \quad (2)$$

The solution is given by:

$$\hat{\vec{x}} = \mathbf{A}^\dagger \vec{y}, \quad \mathbf{A}^\dagger := \mathbf{A}^\top (\mathbf{A}\mathbf{A}^\top)^{-1} \quad (3)$$

This approach is used when direct inversion of the matrix \mathbf{A} is not feasible, a common occurrence in practical scenarios.

Basis Pursuit (BP) method [15] is designed to find the sparsest solution to a linear system by minimizing the l_1 norm:

$$\hat{\vec{x}} = \arg \min_{\vec{x}} \|\vec{x}\|_1 \quad \text{subject to} \quad \vec{y} = \mathbf{A}\vec{x}, \quad (4)$$

To account for noise in the measurements, Basis Pursuit Denoising (BPDN) [16] extends BP by incorporating a regularization term that balances sparsity with fidelity to the observed data:

$$\hat{\vec{x}} = \arg \min_{\vec{x}} \|\vec{x}\|_1 \quad \text{subject to} \quad \|\mathbf{A}\vec{x} - \vec{y}\|_2 \leq \varepsilon, \quad (5)$$

where ε is a tolerance parameter that specifies the acceptable level of error or deviation from the exact solution. It reflects the expected noise level or uncertainty in the measurements. For solving both BP and BPDN problems, the *SPGL1* library is utilized [17]

Another two advanced iterative methods are the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [18] and the Fast Adaptive Shrinkage/Thresholding Algorithm (FASTA) [19]. They effectively solve optimization problems of the following form:

$$\min_{\vec{x}} \left(f(\mathbf{A}\vec{x}) + g(\vec{x}) \right) \quad (6)$$

where $f(\mathbf{A}\vec{x})$ is a smooth function and $g(\vec{x})$ is a possibly non-smooth function, often used to promote sparsity in the solution $\hat{\vec{x}}$. These algorithms can be used to solve the Sparse Least Squares (SLS) problem, which is formulated as:

$$\min_{\vec{x}} \left(\|\mathbf{A}\vec{x} - \vec{y}\|_2^2 + \lambda \|\vec{x}\|_1 \right), \quad (7)$$

where $\|\mathbf{A}\vec{x} - \vec{y}\|_2^2$ represents the smooth function $f(x)$ and $\lambda \|\vec{x}\|_1$ acts as the sparse-promoting function $g(x)$.

D. Coarse-to-Fine Approach (CTF)

The CTF approach is a multi-scale technique used in image processing and signal reconstruction. This strategy involves progressively refining the resolution of the solution. It leverages the idea that solutions at coarser scales can provide good initial estimates for finer scales, thus potentially reducing the computational complexity and improving convergence rates. We have applied the CTF strategy to all the optimization problems defined earlier in this section. For simplicity, let us consider the PI method that solves the linear least square problem presented in (2). We start with a low-resolution matrix at scale k , $\mathbf{A}^{(k)}$, for which the estimated solution of the problem is given by $\hat{\vec{x}}^{(k)} = \mathbf{A}^{(k)\dagger} \vec{y}$. By implementing Otsu's thresholding technique on $\hat{\vec{x}}^{(k)}$, one can determine the support set as follows:

$$T = \tau_{\text{Otsu}} \left(\hat{\vec{x}}^{(k)} \right), \quad \hat{\Omega}^{(k)} = \left\{ i : \hat{x}_i^{(k)} > T \right\} \quad (8)$$

We introduce the operator \mathcal{P} that maps the support set of a low-resolution matrix to the support set of a higher-resolution matrix, where each element of the former is mapped to a subset of the latter. \mathcal{P} thus propagates the support set, $\hat{\Omega}^{(k)}$, to the subsequent, higher resolution level. The propagated support set, $\tilde{\Omega}^{(k+1)}$, which satisfies $n_{k+1} > |\tilde{\Omega}^{(k+1)}| > |\hat{\Omega}^{(k)}|$, is employed with the finer matrix $\mathbf{A}^{(k+1)}$ to refine the solution iteratively. This iterative enhancement is captured in (9) and continues progressively until the finest scale (full-resolution matrix) is reached:

$$\tilde{\Omega}^{(k+1)} = \mathcal{P}(\hat{\Omega}^{(k)}), \quad \hat{\vec{x}}^{(k+1)} = \mathbf{A}_{(\cdot, \tilde{\Omega}^{(k+1)})}^{(k+1)\dagger} \vec{y} \quad (9)$$

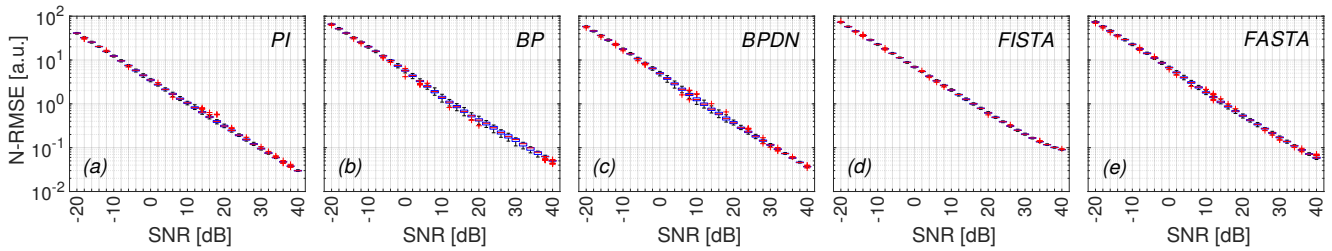


Fig. 2: Robustness to noise: Performance of the Coarse-to-Fine scheme for multiple reconstruction algorithms across varying noise levels, $s_{GT} = 8$

IV. DATA COLLECTION AND ANALYSIS RESULTS

A. Data Acquisition Process

1) *Virtual 3×3 Camera Array*: We sequentially position the camera at nine distinct locations within a plane to emulate a 3×3 camera array. We capture an image at each location and later compile these nine images to create a dataset equivalent to what a 3×3 camera array would capture in a single shot, see Fig. 1 (a).

2) *Virtual Imaging Target*: A virtual cubic volume is defined by systematically moving the THz source across a 3D grid within this volume using a Universal Robot UR5e. The robot ensures precise placement with a delay of 200 ms between movements. This setup captures the system response to point sources located on a uniform grid within the cubic volume.

3) *Data Synthesizing*: At each point within the 3D grid of the virtual cube, our virtual 3×3 camera array captures nine images of the THz source. At each camera position, an image is captured with a 33×33 pixels resolution, resulting in a combined resolution of $9 \times 33 \times 33$ pixels for each grid point. These images are synthesized to simulate the effect of a 3×3 camera array capturing images of the entire cubic volume in one shot.

4) *Empirical Sensing Matrix*: The nine images captured for each grid point are converted into a vector $\vec{y}_i \in \mathbb{R}^m$. The sensing matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is empirically constructed as:

$$\mathbf{A} := [\vec{y}_k]_{k=1}^n \quad (10)$$

5) *Experimental Setup*: The virtual cube used in the experiment has an edge length of 12 cm and a grid step size of 0.8 cm, leading to $n = 16^3$ grid points. The camera array is positioned 20 cm from the cube's center. For all experiments, ground truth (GT) sparse targets, $\vec{x} \in \mathbb{R}^n$, are simulated by activating s_{GT} components chosen uniformly across the cube. Measurements, \vec{y} , follow the model in (1), with statistics derived from 64 realizations of measurement noise, \vec{n} .

B. Stability to Noise Fluctuations

The CTF method enhances stability against noise by focusing computational efforts on columns indexed by $\tilde{\Omega}^{(k+1)}$ at each scale $k + 1$, reducing dimensionality and minimizing noise impact, as demonstrated in Fig. 2. The PI method exhibits remarkable noise stability, with RMSE decreasing from 40 to 0.03 as SNR ranges from -20 dB to 40 dB, highlighting its robustness within a CTF framework (Fig. 2 (a)). Additionally, the CTF framework leads to an RMSE reduction by a factor of approximately 3.8, as shown in Fig. 3 (left). Evaluations of the PI method, both with and without CTF integration, reveal significant improvements, showcasing the advantages of incorporating the CTF framework.

We present visual reconstructions of object "4" using the PI approach, with and without CTF integration. Fig. 3 (right) shows comparisons between the ground truth (GT), the PI method without CTF (PI), and the PI method with CTF integration (PI-CTF), demonstrating the enhanced contrast achieved through the CTF framework. Additionally, the figure also displays reconstructions of object "4" through progressively thicker High-density Polyethylene (HDPE) plates of 2 mm, 4 mm, and 6 mm, showcasing THz radiation's capacity to penetrate materials opaque to visible and infrared light. However, poor SNR degrades image quality beyond 6 mm.

Both BP and BPDN algorithms exhibit similar noise stability, with RMSE decreasing from 60 to about 0.04, as shown in Fig. 2 (b) and (c). Despite being slightly less effective than the PI method, they benefit from the CTF framework, enhancing reconstruction accuracy as noise levels decrease. However, the FISTA and FASTA algorithms show less improvement within the CTF framework (Fig. 2 (d) and (e)), indicating that these iterative methods may not fully leverage the incremental refinements of the CTF framework.

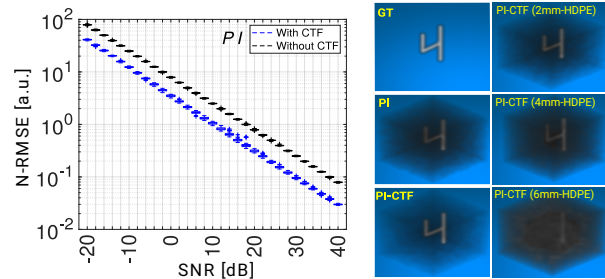


Fig. 3: (left) RMSE vs. SNR for the PI method with/without CTF integration, $s_{GT} = 8$. (right) Ground truth (GT) and reconstructions of object "4" for PI and (PI-CTF) algorithms, followed by PI-CTF through increasing HDPE plate thicknesses of 2mm, 4mm, and 6mm.

C. Stability to Sparsity Levels

Finally, we evaluated the performance of the reconstruction algorithms endowed with a CTF structure for different levels of sparsity, ranging from $s_{GT} = 1$ to $s_{GT} = 256$. Our findings show that the median of the RMSE begins at approximately 0.007 for a sparsity level of $s_{GT} = 1$ across all methods and remains below 0.1 as the sparsity level increases to 8, as illustrated in Fig. 4. However, beyond a sparsity level of $s_{GT} = 8$, the RMSE increases, with its median fluctuating between 0.1 and 0.3, which indicates a rise in reconstruction error as sparsity increases. Average RMSE values per percentiles versus s_{GT} are reported for the CTF-PI method in Table 1 up to $s_{GT} = 16$.

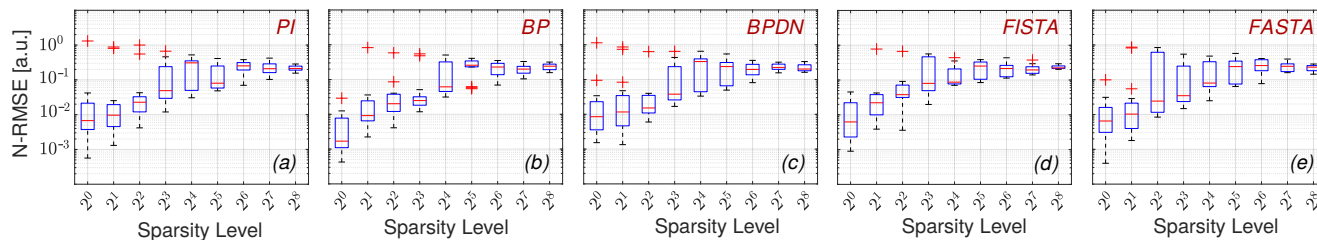


Fig. 4: Robustness to sparsity: Performance of the Coarse-to-Fine scheme for multiple reconstruction algorithms across varying sparsity, SNR = 40.

Target Sparsity	N-RMSE per Percentile [a.u.]			
	0-75%	75-85%	85-95%	95-99%
1	0.0067	0.0227	0.0337	0.6762
2	0.0079	0.0214	0.4151	0.8427
4	0.0173	0.0360	0.2978	0.7767
8	0.0380	0.3088	0.4538	0.5635
16	0.1620	0.3555	0.4092	0.4891

Table 1: Per-percentile average values of normalized RMSE obtained using PI integrated into the CTF framework to reconstruct punctual targets with varying sparsity.

V. CONCLUSION

In this work, we have considered the problem of retrieving the 3D structure of sparse targets in light field THz imaging. To cope with the large ambient dimension with limited computational power without compromising accuracy, a Coarse-to-Fine methodology has been introduced that solves a sequence of sparse reconstruction problems at scales of increasing resolution. We have provided a comprehensive performance evaluation of well-known least-squares and sparse reconstruction algorithms optimized through the CTF approach. Our results underscore the effectiveness of the CTF strategy in enhancing these algorithms' performance. Specifically, our findings highlight the computational advantages of integrating the CTF framework with the PI method, resulting in an RMSE reduction by a factor of approximately 3.8, as illustrated in Fig 3 (left). Additionally, the results demonstrated that the CTF framework has improved the performance of all other algorithms examined in this study, highlighting its broad applicability.

Regarding sparsity, the CTF approach proved efficient across various sparsity levels. Our findings indicate that the RMSE remained consistently low, below 0.1 for sparsity levels up to 8. Fig. 4 shows that integrating the CTF strategy maintained good reconstruction accuracy across all tested methods.

However, the CTF approach is not without its limitations. A significant challenge lies in the potential error propagation through the cascade of reconstruction stages. If the initial coarse reconstructions suffer from inaccuracies, particularly in estimating support, these errors can cascade through subsequent refinements, leading to significant degradation in the final reconstruction quality. Moreover, PI-based approaches are inherently more computationally intensive than transpose-based ones, such as [4]. Additionally, the iterative refinement process can lead to increased memory usage, making it challenging for systems with limited resources.

ACKNOWLEDGMENT

This work received partial funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101019972).

The TWS-ID02 THz camera was provided courtesy of TicWave-Solutions GmbH.

REFERENCES

- [1] P. Hillger, J. Grzyb, R. Jain, and U. R. Pfeiffer, "Terahertz imaging and sensing applications with silicon-based technologies," *IEEE Transactions on Terahertz Science and Technology*, vol. 9, no. 1, pp. 1–19, 2019.
- [2] A. Y. Pawar, D. D. Sonawane, K. B. Erande, and D. V. Derle, "Terahertz technology and its applications," *Drug Invention Today*, vol. 5, no. 2, pp. 157–163, 2013. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0975761913000264>
- [3] R. Jain, J. Grzyb, and U. R. Pfeiffer, "Terahertz light-field imaging," *IEEE Transactions on Terahertz Science and Technology*, vol. 6, no. 5, pp. 649–657, 2016.
- [4] U. R. Pfeiffer and A. Kutaish, "Terahertz light-field imaging with silicon technologies," *IEEE Open Journal of the Solid-State Circuits Society*, vol. 4, pp. 1–11, 2024.
- [5] Y. Jiang, G. Li, H. Ge, F. Wang, L. Li, X. Chen, M. Lu, and Y. Zhang, "Machine learning and application in terahertz technology: A review on achievements and future challenges," *IEEE Access*, vol. 10, pp. 53 761–53 776, 2022.
- [6] S.-J. Wei, X.-L. Zhang, J. Shi, and G. Xiang, "Sparse reconstruction for SAR imaging based on compressed sensing," *Progress in electromagnetics research*, vol. 109, pp. 63–81, 2010.
- [7] G. Zeng, C. Zhang, X. Zhou, Q. Sun, and Y. Xiong, "Research on reconstruction algorithm of terahertz signal based on compressed sensing," in *2022 IEEE 3rd China International Youth Conference on Electrical Engineering (CIYCEE)*, 2022, pp. 1–6.
- [8] C. La and M. N. Do, "Tree-based orthogonal matching pursuit algorithm for signal reconstruction," in *2006 International Conference on Image Processing*, 2006, pp. 1277–1280.
- [9] M. Heredia Conde, K. Hartmann, and O. Loffeld, "Structure and rank awareness for error and data flow reduction in phase-shift-based tof imaging systems using compressive sensing," in *2015 3rd International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*, 2015, pp. 144–148.
- [10] S. L. Bangare, A. Dubal, P. S. Bangare, and S. Patil, "Reviewing Otsu's method for image thresholding," *International Journal of Applied Engineering Research*, vol. 10, no. 9, pp. 21 777–21 783, 2015.
- [11] Y. Kang, K. Engelke, and W. Kalender, "A new accurate and precise 3-D segmentation method for skeletal structures in volumetric CT data," *IEEE Transactions on Medical Imaging*, vol. 22, no. 5, pp. 586–598, 2003.
- [12] M. A. Iqbal, M. Heredia Conde, A. Anghel, and M. Datcu, "Coarse-to-fine estimation: Compressive sensing for high-resolution inverse SAR imaging," in *Proceedings of the European Synthetic Aperture Radar (EuSAR 2024)*, Germany, April 2024.
- [13] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM Journal on Computing*, vol. 24, no. 2, pp. 227–234, 1995. [Online]. Available: <https://doi.org/10.1137/S0097539792240406>
- [14] J. C. A. Barata and M. S. Hussein, "The Moore–Penrose pseudoinverse: A tutorial review of the theory," *Brazilian Journal of Physics*, vol. 42, pp. 146–165, 2012.
- [15] P. Dhole and S. D. Borde, "Sparse signal reconstruction using basis pursuit algorithm," *Signal*, vol. 3, no. 5, 2015.
- [16] R. Baraldi, R. Kumar, and A. Aravkin, "Basis pursuit denoise with nonsmooth constraints," *IEEE Transactions on Signal Processing*, vol. 67, no. 22, pp. 5811–5823, 2019.
- [17] E. van den Berg and M. P. Friedlander, "SPGL1: A solver for large-scale sparse reconstruction," December 2019, <https://friedlander.io/spgl1>.
- [18] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences*, vol. 2, no. 1, pp. 183–202, 2009.
- [19] T. Goldstein, C. Studer, and R. Baraniuk, "A field guide to forward-backward splitting with a FASTA implementation," *arXiv preprint arXiv:1411.3406*, 2014.