University of Wollongong

Research Online

University of Wollongong Thesis Collection 1954-2016

University of Wollongong Thesis Collections

2015

Compressed sensing for enhanced through-the-wall radar imaging

Van Ha Tang
University of Wollongong, vht986@uow.edu.au

Follow this and additional works at: https://ro.uow.edu.au/theses

Recommended Citation

Compressed Sensing for Enhanced Through-the-Wall Radar Imaging

A thesis submitted in partial fulfilment of the requirements

for the award of the degree

Doctor of Philosophy

from

University of Wollongong

by

Van Ha Tang

School of Electrical, Computer and Telecommunications

Engineering

August 2015

Statement of Originality

I, Van Ha Tang, declare that this thesis, submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has not been submitted for qualifications at any other academic institution.

Van Ha Tang

August 2015

Contents

List of Figures

List of Tables

Acronyms

AWG	Additive white Gaussian
BCS	Bayesian compressed sensing
BP	Basis pursuit
BPDN	Basis pursuit denoising
CS	Compressed sensing
DPSS	Discrete prolate spheroidal sequence
EM	Electromagnetic
FAR	False alarm rate
MAP	Maximum a posteriori
OMP	Orthogonal matching pursuit
PCA	Principal component analysis
PD	Probability of detection
PDF	Posterior distribution function
PSNR	Peak signal-to-noise ratio
RCS	Radar cross section
RIP	Restricted isometry property
RMSE	Root mean square error
ROC	Receiver operation characteristics
TCR	Target-to-clutter ratio
TWR	Through-the-wall radar

TWRI   Through-the-wall radar imaging

SAR   Synthetic aperture radar

SBL   Sparse Bayesian learning

SVD   Singular value decomposition

Abstract

Through-the-wall radar imaging (TWRI) is an emerging technology that aims to capture scenes behind walls and other visually opaque materials. The abilities to sense through walls are highly desirable for both military and civil applications, such as search and rescue missions, surveillance, and reconnaissance. TWRI systems, however, face with several challenges including prolonged data acquisition, large objects, strong wall clutter, and shadowing effects, which limit the radar imaging performances and hinder target detection and localization.

Compressed sensing (CS) is a new area of signal processing that has received considerable attention recently. The CS theory enables sparse signals to be reconstructed precisely from far fewer samples than what is required by the Nyquist-Shannon sampling theorem. This dissertation focuses on developing advanced TWRI techniques based on CS and Bayesian CS, an extension of the CS framework. Three different radar imaging approaches are proposed in this project.

The first approach is motivated by the fact that the sparsity assumption of the scene is not usually satisfied due to multipath propagations, wall reflections, or extended objects such as people or furniture. To overcome this problem, we propose a two-stage CS-based TWR image formation algorithm. It incorporates a dictionary constructing from the complex wavelet Gabor basis used to sparsely represent the scene image. Incorporating the sparsifying dictionary increases the sparsity and directional selectivity in space-frequency domain, thereby leading to

more effective CS techniques.

The second approach is developed for addressing the scene reconstruction problem in the presence of wall clutter under CS context. Although not all same frequency measurements are available at each antenna location, the proposed approach estimates the antenna signal coefficients simultaneously by exploiting both the sparsity and the inter-signal correlations among antenna signals. A joint Bayesian sparse model is employed to reconstruct the antenna signal coefficients and to estimate the image of the scene. For scene reconstruction, a compact linear model is developed, whereby both the measurement vector and the dictionary are compressed, leading to a more efficient Bayesian scene reconstruction. Furthermore, a subspace-projection technique is applied directly to the recovered antenna signal coefficients to suppress wall clutter and enhance image quality and target detection.

In the third approach, a joint Bayesian sparse representation is proposed for compressed multi-channel TWRI where the same scene is sensed from multiple views or multiple polarizations. The proposed approach combines reduced measurement sets collected from different channels for enhancing TWR imaging. The antenna signal coefficients associated with different channels are first simultaneously estimated. Then, a subspace-projection technique is used for removing the wall returns. For multi-channel image reconstruction, a multi-task imaging model comprising a composite measurement vector and dictionary is formulated, and then the scene images related to individual channels and a composite image of the scene are simultaneously reconstructed using joint Bayesian CS, taking inter-channel dependencies into the imaging model.

All the proposed approaches are evaluated using simulated and real radar data. The experimental results prove the efficiency in data acquisition, image formation, target detection and localization. Using a small percentage of the full data volume, the proposed approaches produce high-resolution and high-quality images that are suitable for indoor target identification and classification.

Acknowledgments

Chapter 1

Introduction

Chapter contents

1.1 Research objectives

Over the past decade, through-the-wall radar imaging has witnessed increasing research interests due to its abilities to image behind-wall and indoor scenes. These scenes are typically inaccessible via optical, acoustical, or thermal sensing. TWRI can be used to determine the building layouts, monitor humans and other targets, and recognize activities inside the building. The ability to see through walls is desirable in numerous applications, from locating hidden hostages in a police operation, tracking concealed hostile forces in a military mission, to locating buried victims of an earthquake or fire [1, 2, 3, 4, 5, 6]. In practical operations, however, TWRI and urban sensing face with many technical challenges including prolonged data collection, strong wall clutter, and shadowing effects. Thus, there is a need to tackle such difficulties and achieve a reliable performance for urban radar sensing systems.

The aim of this research is to develop radar imaging techniques for detecting and localizing targets behind walls and opaque materials using a small percentage of the entire data volume. Towards this goal, novel algorithms based on compressed sensing (CS) and Bayesian compressed sensing are proposed to enhance image quality and resolution, improve speed of operation, and reduce the cost and time of data acquisition and processing. Several CS-based approaches have been proposed for TWRI, but this research investigates the problem in which the potential of CS has not been fully examined. The proposed approaches intend to relax constraints on the sparsity assumption of the scene, signal sampling schemes, and logistic difficulties in data acquisition. The imaging problems are addressed in compressed TWR sensing operations that only a reduced set of antenna locations and a subset of frequencies are used for sensing. This research is important because in many practical applications, the full measurements are impossible to obtain and several measurements or samples are lost due to frequency interferences and radar jamming.

The problem of scene reconstruction involving extended targets is addressed in this research. The point-like target assumption is overcome by incorporating a sparsifying dictionary to represent the scene image. Moreover, we consider the stationary target detection and localization in the presence of the wall returns. The problem of image formation and wall clutter rejection is handled by employing a Bayesian CS framework. Furthermore, this research aims to address the problem of compressed multi-channel TWRI. Multi-view, multi-location or multi-channel TWRI improves stationary target identification by combining multiple data sets acquired from different sensing locations or channels. Collecting data at several sensing channels enhances imaging visibility, but also leads to prolonged data acquisition time, complex computation, and expensive hardware. Imaging algorithms based on joint Bayesian compressed sensing are proposed in this research to take the inter-channel dependencies into account and thereby increasing the capability of target detection, localization, and identification.

1.2   Research contributions

The proposed research has yielded solutions that improve the capabilities of TWRI systems. A scene behind-the-wall and inside enclosed structures can be sensed at a faster speed with far fewer measurements and thereby reducing the time of data acquisition and cost of operation. The main contributions of this research project are in the area of the applications of CS and Bayesian CS to TWRI, and are outlined as follows.

The first contribution of this dissertation involves investigating of commonly used sparsifying basis, such as wavelet, Gabor for increasing the scene sparsity. Motivated by the fact that the performance of CS relies on the sparsity of the underlying signal, in TWR applications, the sparsity assumption of the scene, however, is violated due to multipath propagations, wall reflections, or extended objects such as people or furniture. Therefore, a two-stage TWR image formation algorithm using CS is proposed. In the first stage, an additional sparsifying dictionary constructed from complex Gabor functions is incorporated in the imaging model for sparse scene representation, which allows CS to be applied for recovering a full data volume from a reduced measurement set. In the second stage, conventional backprojection methods, such as delay-and-sum beamforming are employed to form an image of the scene. Experiential results on both synthetic and real data show that the proposed approach forms a high-quality image of the scene in terms of target-to-clutter ratio, especially when the measurements are drastically reduced.

The second contribution is to develop an imaging model for scene reconstruction in conjunction with wall clutter rejection. In TWRI, the front wall returns usually dominate the target reflections, rendering the target detection difficult, or even impossible. Several CS-based TWRI approaches assume that the wall clutter and reverberations have been removed prior to applying CS by having access to the background or reference scene. These approaches are effective in

removing wall radar signals and applicable for applying CS techniques, but impossible in practical applications. A sparse Bayesian approach is proposed in this research to address the problem of wall clutter mitigation and image reconstruction under CS context. In the proposed approach, a joint Bayesian sparse approximation is employed to jointly reconstruct the antenna signal coefficients, by exploiting the signal sparsity and inter-signal correlations among received signals. This is in contrast to existing single-signal CS model that recovers antenna signals independently, considering the intra-signal sparsity structure only. Furthermore, a subspace-projection technique is applied directly to the recovered signal coefficients to remove those associated with the wall returns. Finally, a compact linear imaging model is formulated for efficient sparse Bayesian scene reconstruction. Experimental results on simulated electromagnetic (EM) data and real radar signals show that the proposed approach enhances TWRI in terms of signal reconstruction accuracy, target-to-clutter ratio, and target detection rate. Especially, when the measurements are drastically reduced or the sampling rates are low, while the existing CS-based imaging models fail to localize the targets, the proposed approach manages to detect the objects behind-the-wall.

Addressing the problem of multi-channel TWRI using joint Bayesian compressed sensing forms the third contribution of this research. Here, the scene is illuminated from multiple vantage points or from a single viewing angle, but with different polarizations. Multi-view, multi-location, or multi-polarization TWR imaging enhances stationary target detection and localization by combining multiple data sets acquired from different sensing locations or polarimetric channels. To date, most of the existing CS-based TWR imaging methods have been proposed for single-channel operation mode where a reduced data set collected from one channel is used to form the scene image. The problem of multi-channel TWR scene reconstruction, incorporating wall-clutter mitigation, under a CS context is addressed in this research. The proposed approach considers the sensing problem where the same scene is sensed by deploying the antenna aperture along

the front and side walls of an enclosed structure or different polarimetric channels. At each channel, different reduced sets of frequencies are used at each antenna. In the proposed approach, the antenna signal coefficients are first jointly reconstructed by a joint Bayesian sparse approximation framework, followed with wall clutter removal using a subspace-projection technique. Finally, a multi-task linear imaging model is developed to combine target coefficients from different channels for jointly reconstructing a composite scene image and images related to different sensing channels. In this model, the fusion is performed in the data level, and the inter-channel correlations are exploited using joint Bayesian sparse learning framework.

1.3 Thesis structure

The thesis is organized as follows:

- **Chapter 1** presents the research project, its objectives, research contributions, and a summary of related publications.

- **Chapter 2** reviews the theory of compressed sensing, a novel sensing paradigm that goes beyond the common Nyquist-Shannon's theory in data acquisition and signal reconstruction.

- **Chapter 3** introduces high-resolution through-the-wall radar imaging and TWR image formation using a backprojection algorithm. This chapter also presents two major compressed sensing techniques for TWRI: CS-based image formation and signal reconstruction. It is followed by discussions highlighting the research gaps that are addressed in this project.

- **Chapter 4** describes a proposed two-stage through-the-wall radar image formation algorithm using CS. The proposed approach incorporates a sparsifying dictionary for scene sparse representation which enhances the applicability of CS to TWR imaging.

- **Chapter 5** proposes a Bayesian scene reconstruction for compressed TWR sensing. The Bayesian sparse framework is employed to model both the intra-signal structure sparsity and inter-signal correlations for jointly reconstructing all the antenna signals, followed with wall clutter suppression using a subspace-projection technique. A compact linear model is formulated for efficient Bayesian scene reconstruction.

- **Chapter 6** presents a joint Bayesian CS model for multi-channel TWRI, where the scene is illuminated from a single vantage point, but with different polarimetric channels or when the same scene is imaged from multiple viewing angles. Given multi-channel compressed data sets, scene images associated with different channels and a composite scene image are simultaneously reconstructed using joint Bayesian CS framework, taking the inter-channel dependencies into the imaging model.

- **Chapter 7** summarizes the research findings and provides concluding remarks and possible future directions.

1.4 Publications

Following is the list of publications arising from this PhD research project, which was conducted from September 2011 to August 2015.

- V. H. Tang, A. Bouzerdoum, and S. L. Phung, "Two-stage through-the-wall radar image formation using compressive sensing," *SPIE Journal of Electronic Imaging*, vol. 22, no. 2, pp. 021 006–1–021 006–10, 2013.

- V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Multi-view indoor scene reconstruction from compressed through-wall radar measurements using a joint Bayesian sparse representation," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Brisbane, Australia, April 2015, pp. 2419–2423.

- V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Enhanced wall clutter mitigation for through-the-wall radar imaging using joint Bayesian sparse signal recovery," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Florence, Italy, May 2014, pp. 7804–7808.

- V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Multi-view TWRI scene reconstruction using a joint Bayesian sparse approximation model," in *Proc. SPIE Defence, Security and Sensing: Compressive Sensing*, vol. 9484, Maryland, USA, April 2015, pp. 9484–1–9484–12.

- V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Enhanced through-the-wall radar imaging using Bayesian compressive sensing," in *Proc. SPIE Defence, Security and Sensing: Compressive Sensing*, vol. 8717, Maryland, USA, May 2013, pp. 87 170I–1–87 170I–12.

- A. Bouzerdoum, F. H. C. Tivive, and V. H. Tang, "Multi-polarization through-the-wall radar imaging using joint Bayesian compressed sensing," in *Proc. IEEE Int. Conf. Digital Signal Processing*, Hong Kong, China, August 2014, pp. 783–788.

- F. H. C. Tivive, A. Bouzerdoum, and V. H. Tang, "Multi-stage compressed sensing and wall clutter mitigation for through-the-wall radar image formation," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, Coruna, Spain, June 2014, pp. 489–492.

Chapter 2

Compressed sensing: A review

Chapter contents

Compressed sensing also known as compressive sensing, or compressive sampling is a recent signal processing technique for efficiently acquiring and reconstructing sparse signals by solving underdetermined linear systems. The breakthrough in CS is that sparse signals can be exactly recovered from far fewer samples than what is required by the conventional Nyquist-Shannon sampling theory [14, 15, 16, 17, 18, 19]. Conventional approaches to sampling signals follow the Nyquist-Shannon's theorem [20, 21] which states that the sampling rate must be at least twice the maximum frequency present in the signals. In fact, this

principle underlies nearly all signal acquisition methods used in practical applications. However, in several emerging applications, the resulting Nyquist rate is so high that the sensing system acquires far many samples, leading to technical challenges in data acquisition, storage, and processing. In some sensing applications, it is even physically impossible to build devices for acquiring samples at such high sampling rate [22]. Thus, CS is considered as a new framework that enables a potentially large reduction in the sampling and computation costs for signal acquisition, representation, and reconstruction.

Compressed sensing theory suggests that it is possible to sense sparse signals by taking far fewer measurements, hence the name compressed sensing [23, 24, 25]. It differs from classical sampling theory in three important aspects. First, traditional sampling theory typically considers infinite-length, continuous-time signals. In contrast, CS is a mathematical framework that focuses on working with finite-dimensional vectors. Second, rather than sampling the signal at specific points in time, CS systems generally collect measurements by using inner products between the signal and more general test functions. Third, the two frameworks differ in the signal reconstruction. In the Nyquist-Shannon framework, signal recovery is performed through *sinc* interpolation. By contrast, in CS, signal recovery is typically achieved using highly non-linear methods [26]. The recovery is enabled through two principles: sparsity and incoherence. The sparsity reflects the fact that many nature signals are sparse or compressible when represented in the proper basis. The incoherence is represented through the isometric property, which is a sufficient condition for sparse signal recovery. In this chapter, we review the principles and fundamental premises underlying CS framework including sparsity, incoherence, conditions for accurate signal recovery, and sparse signal recovery algorithms.

2.1 Signal model

In this section, we present the standard finite-dimensional CS signal model. Assume that we obtain a collection of N samples of a signal arranged in a column vector $\mathbf{z}$. In cases where these samples are acquired from an image or other multidimensional signals, the data samples can be lexicographically stacked into the column vector. The vector $\mathbf{z}$ is modeled as the superposition of the clean signal $\bar{\mathbf{z}}$ and noise term $\mathbf{v}$:

$$\mathbf{z} = \bar{\mathbf{z}} + \mathbf{v}. \tag{2.1}$$

In CS, instead of acquiring the full N data samples, we collect a far fewer K number of linear measurements of $\mathbf{z}$. These recorded values are arranged into a $K \times 1$ measurement vector $\mathbf{y}$. We can represent the acquisition process mathematically as

$$\mathbf{y} = \mathbf{\Phi}\,\mathbf{z} = \mathbf{\Phi}\,\bar{\mathbf{z}} + \bar{\mathbf{v}}, \tag{2.2}$$

where $\mathbf{\Phi}$ is a $K \times N$ sensing matrix and $\bar{\mathbf{v}} = \mathbf{\Phi}\,\mathbf{v}$. Typically in CS, $\mathbf{\Phi}$ is designed with some element of randomness, and depends on specific applications. For example, in stepped-frequency TWRI, $\mathbf{\Phi}$ is known as selection matrix, which indicates the subset of antennas and the frequency bins used.

CS aims to reconstruct the clean signal $\bar{\mathbf{z}}$ using the compressed measurement vector $\mathbf{y}$. For accurate reconstruction, prior knowledge about the signal and the structure model of sensing matrix $\mathbf{\Phi}$ should be exploited since the linear model in Eq. (2.2) is underdetermined ($K < N$) and therefore has an infinite number of candidate solutions. In the next sections, we review the underlying principles in CS as well as the conditions for accurate signal recovery.

2.2 Sparsity and compressibility

Sparse or compressible signals can be defined due to the fact that there are many high-dimensional signals having concise representations when expressed in a

convenient basis or dictionary. These bases can be Fourier, wavelet, Gabor or other basis functions that promote sparse representations [27, 28, 29]. The signals can be represented by using relatively few large signal coefficients as such large coefficients hold most of the signal energy.

Mathematically, let Ψ denote an $N \times N$ basis matrix containing in its columns the basis functions: $\Psi = [\Psi_0, \Psi_1, \dots, \Psi_{N-1}]$. The signal $\bar{z}$ can be expanded as

$$\bar{z} = \sum_{i=0}^{N-1} x_i \Psi_i, \tag{2.3}$$

or in vector-matrix form as

$$\bar{z} = \Psi x \tag{2.4}$$

where $x = [x_0, x_1, \dots, x_{N-1}]^T$ is a vector of size $N \times 1$ containing the signal coefficients. The coefficient vector x is said to be S-sparse if it contains just S nonzero entries ($S \ll N$). The ℓ_0-norm can be used to count the number of nonzero elements in x: $\|x\|_0 = S$. We refer to the set of positions of the nonzero entries of x as the *support* of x. The signal expansion in Eq. (2.4) is also known as sparse representation.

As an example, we introduce here the signal representation in the TWRI application. This sensing operations transmit and receive signals at several antenna locations to image a scene behind walls. Figure 2.1(a) illustrates the frequency signal z received at one antenna location. This signal contains 801 frequency samples. Obviously the signal itself is not sparse: the ℓ_0-norm value is $\|z\|_0 = 801$. Figure 2.1(b) depicts the coefficient vector x obtained by representing z using the Fourier basis. We can observe from the figure that the signal coefficient vector is relatively sparse in the transformed domain. The ℓ_0-norm value is $\|x\|_0 = 17$.

Sparsity implies that when a signal has a sparse representation, the small signal coefficients can be discarded without much perceptual loss. Let $\bar{z}_S$ denote the signal obtained by keeping only the samples corresponding to the S largest
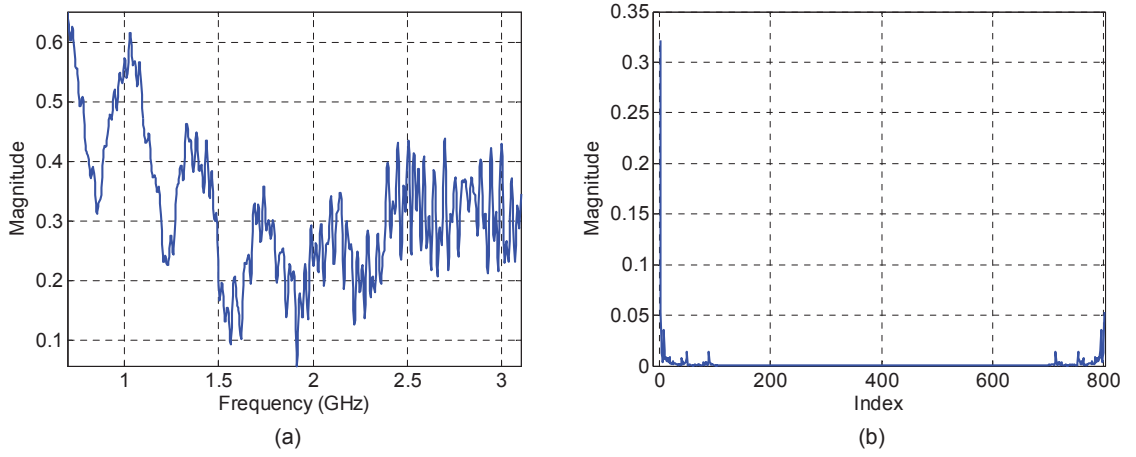
Figure 2.1: The sparse representation for a signal: (a) frequency TWR signal $\mathbf{z}$ received at one antenna ($\|\mathbf{z}\|_0 = 801$); (b) the signal coefficient vector $\mathbf{x}$ represented using the Fourier dictionary ($\|\mathbf{x}\|_0 = 17$).

values of the coefficient vector $\mathbf{x}$:

$$\bar{\mathbf{z}}_S = \mathbf{\Psi}\,\mathbf{x}_S, \tag{2.5}$$

where $\mathbf{x}_S$ denotes the nearest S-sparse vector to vector $\mathbf{x}$. Assuming that the sparse basis $\mathbf{\Psi}$ is orthonormal, we have $\|\mathbf{z} - \mathbf{z}_S\|_2 = \|\mathbf{x} - \mathbf{x}_S\|_2$. If $\mathbf{x}$ is sparse and the sorted magnitudes of the x_i decay quickly, then $\mathbf{x}$ is well approximated by $\mathbf{x}_S$ and the error $\|\mathbf{z} - \mathbf{z}_S\|_2$ is very small. In other words, we can discard a large fraction of the insignificant coefficients without much loss of signal information.

To illustrate the efficiency of sparse representation, we use here the TWR image sparse representation. Figures 2.2(a)–(b) show, respectively, an original scene image and its corresponding signal vector $\mathbf{z}$ obtained by stacking the image into a column vector. It is observed that the original signal here is not sparse. Figures 2.2(c)–(d) present, respectively, the image reconstructed by keeping only 5% of the significant Gabor coefficients and the Gabor coefficient vector $\mathbf{x}$. It is observed that the reconstructed image is visually similar to the original image. This example illustrates that if a natural signal is compressible or approximately sparse in some transformed domain, it can be compressed efficiently while maintaining the salient information.
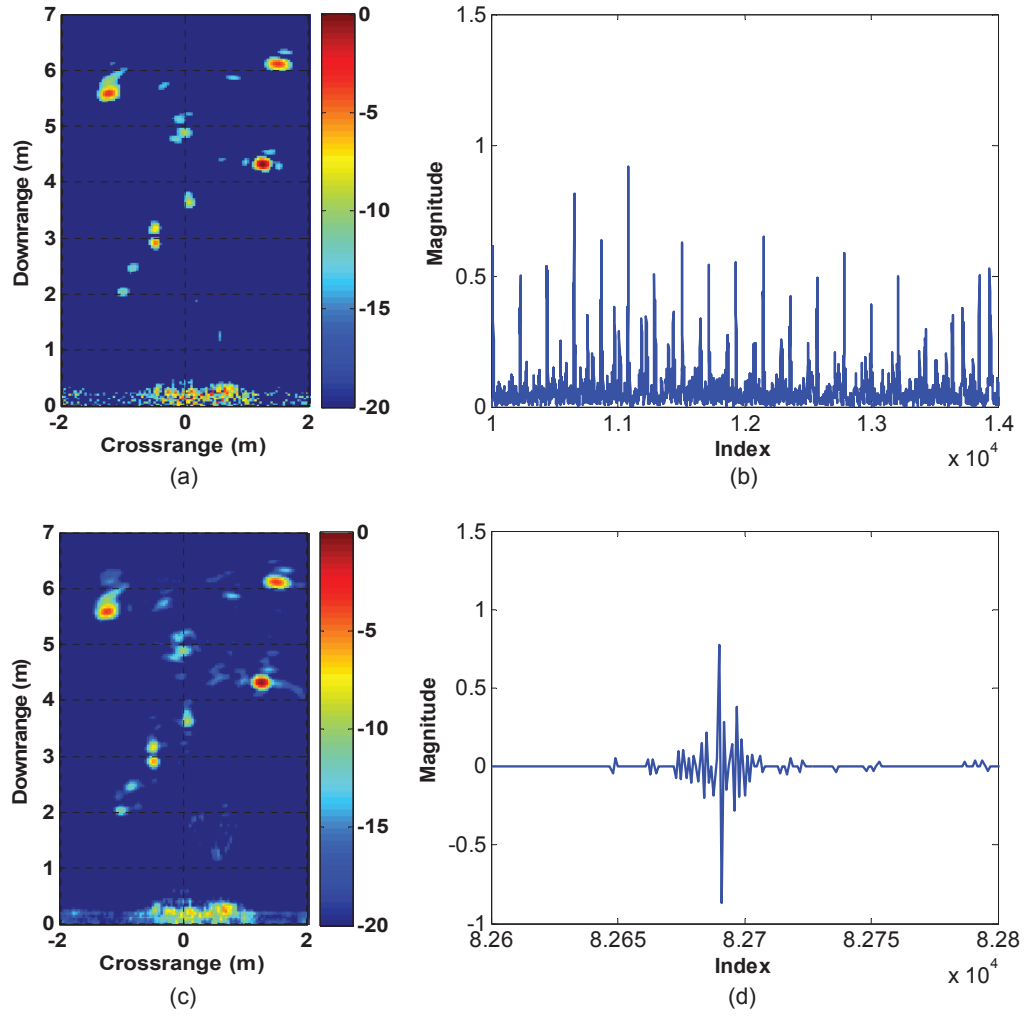
Figure 2.2: The sparse representation for a TWR scene image: (a) original TWR scene image; (b) the image stacked into the full signal **z**; (c) reconstructed image using 5% large Gabor coefficients; (d) the coefficient vector **x** represented using Gabor basis.

The insights into signal sparse representation can be discussed by using the ℓ_1-norm and ℓ_2-norm. The ℓ_1 norm measures the absolute sum of the entries of vector **x**:

$$\|\mathbf{x}\|_1 = \sum_{i=0}^{N-1} |x_i|, \tag{2.6}$$

and ℓ_2 norm measures the sum of squared magnitudes of the elements of **x**:

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=0}^{N-1} |x_i|^2}. \tag{2.7}$$

Unlike the ℓ_0 measure, the ℓ_1 and ℓ_2 norms have the formal mathematical defini-

13

tion of a norm [30], and both are convex functions of **x** [31]. The ℓ_1 norm is directly related to sparsity: it tends to be small for sparse signals. Therefore, instead of using ℓ_0-norm, ℓ_1 norm is used to enforce the sparsity of signals in formulating optimization problems for signal reconstruction [32]. In certain real applications, the sparsity structure of signals tends to be grouped. That is the non-zero coefficients are grouped together. In this case, the ℓ_2-norm is used in combination with the ℓ_1-norm to enforce the group sparsity structure for signal recovery model. Sparsity and compressibility play an important role in the acquisition process and determine how efficiently one can acquire signals nonadaptively [14, 16, 23].

2.3 Sparse signal recovery and guarantees

We now address the problem of how a sparse model can be used to recover a signal **z** from a compressed measurement vector **y**. From Eqs. (2.2) and (2.4), we have

$$\mathbf{y} = \mathbf{\Phi}\,\bar{\mathbf{z}} + \bar{\mathbf{v}} = \mathbf{\Phi}\,\mathbf{\Psi}\,\mathbf{x} + \bar{\mathbf{v}} = \mathbf{D}\,\mathbf{x} + \bar{\mathbf{v}}, \tag{2.8}$$

where $\mathbf{D} = \mathbf{\Phi}\,\mathbf{\Psi}$ is a $K \times N$ matrix, **x** is a S-sparse or compressible vector of length N. Most CS recovery algorithms can be used for approximating a sparse vector $\hat{\mathbf{x}}$ that satisfies $\mathbf{y} \approx \mathbf{D}\,\hat{\mathbf{x}}$. Once the sparse coefficient vector has been approximated, an estimated signal can be obtained as $\hat{\mathbf{z}} = \mathbf{\Psi}\,\hat{\mathbf{x}}$.

2.3.1 Recovery guarantees

Performance guarantees for several CS recovery algorithms depend on the sensing matrix $\mathbf{\Phi}$ and the sparsity basis $\mathbf{\Psi}$. In other words, the recovery is possible with high probability if the restricted isometry property (RIP) holds for the the sensing and sparsity matrices [17, 33]:

Definition 1 *An $K \times N$ measurement matrix $\mathbf{\Phi}$ is said to satisfy the RIP of order S with respect to the $N \times N$ orthonormal basis $\mathbf{\Psi}$ if there exists a constant $\sigma_S \in (0,1)$ such that*

$$(1 - \sigma_S)\|\mathbf{x}\|_2^2 \leq \|\mathbf{\Phi}\,\mathbf{\Psi}\,\mathbf{x}\|_2^2 \leq (1 + \sigma_S)\|\mathbf{x}\|_2^2 \tag{2.9}$$

holds for all sparse vectors $\mathbf{x}$ *with* $\|\mathbf{x}\|_0 \leq S$. *The parameter* σ_S *is known as the isometric constant of order S.*

Through the properties of the dictionary $\mathbf{D}$, the condition (2.9) is equivalent to

$$(1 - \sigma_S)\|\mathbf{x}\|_2^2 \leq \|\mathbf{D}\,\mathbf{x}\|_2^2 \leq (1 + \sigma_S)\|\mathbf{x}\|_2^2, \tag{2.10}$$

holding for all vectors $\mathbf{x}$ with $\|\mathbf{x}\|_0 \leq S$. Since this requirement depends only on the dictionary $\mathbf{D}$, we refer to this condition simply as $\mathbf{D}$ satisfying the RIP.

Another condition that CS recovery algorithms rely on is the coherence between the sparsity basis $\boldsymbol{\Psi}$ and sensing matrix $\boldsymbol{\Phi}$. The coherence measures the largest correlation between any two columns of $\boldsymbol{\Phi}$ and $\boldsymbol{\Psi}$ [16, 34, 35].

Definition 2 *The coherence between the sensing matrix* $\boldsymbol{\Phi}$ *and the sparsity basis* $\boldsymbol{\Psi}$ *is*

$$\mu(\boldsymbol{\Phi}, \boldsymbol{\Psi}) = \sqrt{N} \max_{0 \leq i,j \leq N-1} \left| \langle \phi_i, \psi_j \rangle \right|. \tag{2.11}$$

If the sensing matrix and sparsity basis contain correlated elements, then the coherence is high. Otherwise, it is small. Generally, the coherence $\mu(\boldsymbol{\Phi}, \boldsymbol{\Psi}) \in \left[1, \sqrt{N}\right]$. For efficient CS recovery, the sensing matrix $\boldsymbol{\Phi}$ should be sufficiently different from the sparsity basis $\boldsymbol{\Psi}$, and thus CS is mainly concerned with low coherence pairs.

The coherence can be used to determine the minimal number of measurements K for stable signal recovery as [16]

$$K \geq C\,\mu(\boldsymbol{\Phi}, \boldsymbol{\Psi})^2\,S\,\log N, \tag{2.12}$$

for some positive constant C. From Eq. (2.12), we can see that the smaller the coherence, the fewer samples are needed. Moreover, it is possible to precisely recover the original signal by measuring just about any set of coefficients which may far less than the signal size. If $\mu(\boldsymbol{\Phi}, \boldsymbol{\Psi})$ is equal or close to one, then we just

need $S \log N$ samples instead of N.

2.3.2 Numerical recovery algorithms

If the conditions for CS performance are satisfied, the S-sparse coefficient vector $\mathbf{x}$ can be recovered from the measurements $\mathbf{y}$. In the noiseless case, the sparse vector $\mathbf{x}$ is obtained by solving the following optimization problem

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{subject to} \quad \mathbf{y} = \mathbf{D}\,\mathbf{x}. \tag{2.13}$$

Unfortunately, this problem is known to be non-deterministic polynomial-time hard (NP-hard) [36, 37]. As discussed above, the ℓ_1 tends to be small for sparse signals and thus it promotes the signal sparsity. Therefore, the ℓ_1-norm is used to enforce the sparsity as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{D}\,\mathbf{x}. \tag{2.14}$$

The optimization problem in (2.14) is generally known as *basis pursuit* (BP) [14, 38]. The BP problem can be cast as a linear program or a second-order cone program. Standard techniques from convex optimization, such as the simplex method or interior point methods, can be used for solving these problems [31]. Popular software packages such as ℓ_1-Magic [39] were developed to solve a particular set of sparse recovery problems. Furthermore, efficient methods including homotopy algorithm [40, 41], spectral projected-gradient algorithm (SPGL1) [42], and alternating direction algorithm (YALL1) [43] are available for solving the BP problem.

In the presence of noise, the equality constraint in (2.14) is recast into a quadratically-constrained ℓ_1 minimization as

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{D}\,\mathbf{x}\|_2 \le \epsilon, \tag{2.15}$$

where ϵ is a noise bound. This optimization problem is known as *basis pursuit denoising* (BPDN) [15, 38, 44]. Standard techniques from convex optimization can be used to solve the BPDN problem [31]. Moreover, several other methods have been designed to solve this problem efficiently including NESTA [45] and SPAMS [46]. Here, it is important to note that the recovery accuracy depends on the selection of the regularization parameter ϵ. Fortunately, techniques based on cross validation have been proposed for estimating the parameter from the compressed measurements [47, 48].

Alternative algorithms to convex optimization-based reconstruction are greedy techniques including orthogonal matching pursuit (OMP) [49, 50] and compressive sampling matching pursuit (CoSaMP) [51]. OMP method attempts to correct the errors in the estimation caused by minor correlations among the columns of the matrix **D**. Specifically, OMP relies on an iterative process to identify the support of **x** one element at a time. At each iteration, a residual vector is correlated against the columns of **D**. The position with the largest inner product is added to the support estimate. Candidate values for the entries of **x** on this support are then computed using least-squared technique, and the residual vector is updated. CoSaMP algorithm can be regarded as a refinement of the OMP idea. The same as OMP, CoSaMP aims to identify a sparse support set that leads to a small residual. However, CoSaMP constructs the support set multiple elements at a time, and at each iteration, elements can be both added to and removed from the estimated support set. Since both OMP and CoSaMP greedy algorithms attempt to estimate the support set, they generally require the sparsity level S of **x** as a stopping criteria. Moreover, for accurate recovery, greedy methods need a larger number of measurements, compared with standard convex optimization algorithms.

2.4 Bayesian compressive sensing

As an extension of standard compressed sensing, a Bayesian compressive sensing (BCS) framework has recently been introduced for estimating sparse signals. The

BCS approach casts the sparse estimation model (2.8) into a Bayesian framework and solves it based on a sparse Bayesian learning technique [52, 53, 54, 55, 56]. The estimation of the sparse vector $\mathbf{x}$ is provided as a full posterior distribution function (PDF), rather than a point estimate as in conventional CS methods.

In this section, we present the BCS framework for formulating and solving the CS inversion problem. Before presenting the CS problem from a Bayesian viewpoint, we note that the CS problem (2.15) can be rewritten equivalently as [55, 56]

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \left\{ \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1 \right\}, \tag{2.16}$$

where λ is a small positive constant.

2.4.1 CS from a Bayesian perspective

Generally, the BCS approach is to find the most likely sparse vector $\mathbf{x}$ assuming that this vector is approximately sparse and the measurement vector $\mathbf{y}$ is corrupted by noise with known distribution. The sparsity is guaranteed by a prior defined on vector $\mathbf{x}$, while the noise is modeled through the likelihood term. Suppose that the noise vector is zero-mean Gaussian with independent and identically distributed (i.i.d) components by variance σ^2 (noise precision $\beta = 1/\sigma^2$). Then, the likelihood of $\mathbf{y}$ is modeled as a multivariate Gaussian function

$$p(\mathbf{y}|\mathbf{x}) = (2\pi\sigma^2)^{-K/2} \exp(-\frac{1}{2\sigma^2} \left\|\mathbf{y} - \mathbf{D}\mathbf{x}\right\|_2^2). \tag{2.17}$$

The sparsity of $\mathbf{x}$ is imposed using a sparsity-promoting prior. Here, a common prior is the separable Laplacian density function

$$p(\mathbf{x}) = (\lambda/2)^N \exp(-\lambda \sum_{i=0}^{N-1} |x_i|). \tag{2.18}$$

By Bayes rule, the posterior for the signal coefficients can be related to the

likelihood and the prior as

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})\, p(\mathbf{x})}{p(\mathbf{y})}. \tag{2.19}$$

We find the signal that maximizes this posterior probability using maximum a posteriori (MAP) estimation. Because the denominator of (2.19) is independent of $\mathbf{x}$, the MAP estimate can be performed by minimizing the negative of the logarithm of the numerator as

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{D}\,\mathbf{x}\|_2^2 + 2\sigma^2\lambda\|\mathbf{x}\|_1. \tag{2.20}$$

The expression in (2.20) is very similar to the minimization problem in (2.16). Therefore, we can consider the conventional CS recovery algorithms as MAP estimates with a prior on the sparse coefficient vector. It is also possible to recast many CS algorithms as MAP estimators with respect to some priors [57, 58].

2.4.2 Sparse Bayesian learning

As presented in Subsection 2.4.1, the conventional CS recovery problem can be regarded to a MAP approximation in a Bayesian linear analysis, with a Laplace sparseness prior on $\mathbf{x}$. To provide a full posterior estimate of $\mathbf{x}$, the Bayesian analysis needs to be performed further, which requires the prior and likelihood to be conjugate. Since the Laplace prior is not conjugate to the Gaussian likelihood and therefore the Bayesian inference can not be performed in closed form [59, 60]. This issue can be addressed by employing the sparse Bayesian learning (SBL) framework. In SBL, instead of using a Laplace prior, the sparsity of $\mathbf{x}$ is enforced using a Gaussian prior:

$$\begin{aligned} p(\mathbf{x}|\boldsymbol{\alpha}) &= \prod_{i=0}^{N-1} \mathcal{N}(x_i \,|\, 0, \alpha_i^{-1}), \\ &= (2\pi)^{-N/2} |\mathbf{A}|^{1/2} \exp\left(-\frac{1}{2}\mathbf{x}^T \mathbf{A}\,\mathbf{x}\right), \end{aligned} \tag{2.21}$$

where $\mathbf{A} = \text{diag}(\alpha_0, \alpha_1, \ldots, \alpha_{N-1})$. Here, it is important to note that the hyperparameters $\boldsymbol{\alpha} = [\alpha_0, \ldots, \alpha_{N-1}]$ are used to control the variance of the Gaussian density function $\mathcal{N}(x_i|0, \alpha_i^{-1})$. During sparse Bayesian learning, most of α_i, for $i = 0, \ldots, N-1$, tend to diverge to infinity, and hence the values x_i are zero due to the zero-mean and zero-variance Gaussian prior at this location. In other words, this prior promotes sparsity for $\mathbf{x}$ [61].

Combining likelihood (2.17) and prior (2.21), by Bayes rule, the posterior for $\mathbf{x}$ can be expressed as

$$p(\mathbf{x}|\mathbf{y}, \boldsymbol{\alpha}, \beta) = \frac{p(\mathbf{y}|\mathbf{x}, \beta)\, p(\mathbf{x}|\boldsymbol{\alpha})}{p(\mathbf{y}|\boldsymbol{\alpha}, \beta)}, \tag{2.22}$$

$$= (2\pi)^{-(K+1)/2} |\boldsymbol{\Sigma}|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\}, \tag{2.23}$$

which is a multivariate Gaussian distribution, $\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with mean and covariance given by [62]

$$\boldsymbol{\mu} = \beta\, \boldsymbol{\Sigma}\, \mathbf{D}^T \mathbf{y}, \tag{2.24}$$

$$\boldsymbol{\Sigma} = (\beta\, \mathbf{D}^T \mathbf{D} + \mathbf{A})^{-1}. \tag{2.25}$$

The learning problem thus becomes searching for the hyperparameters $\boldsymbol{\alpha}$ and β using a type-II ML procedure. In this type-II ML technique, the hyperparameters $\boldsymbol{\alpha}$ and β are found by maximizing the marginal likelihood, or equivalently, its logarithm $\mathcal{L}(\boldsymbol{\alpha}, \beta)$:

$$\mathcal{L}(\boldsymbol{\alpha}, \beta) = \log p(\mathbf{y}|\boldsymbol{\alpha}, \beta),$$

$$= \log \int p(\mathbf{y}|\mathbf{x}, \beta)\, p(\mathbf{x}|\boldsymbol{\alpha})\, d\mathbf{x},$$

$$= -\frac{1}{2}[K \log 2\pi + \log |\mathbf{C}| + \mathbf{y}^T \mathbf{C}^{-1} \mathbf{y}], \tag{2.26}$$

where $\mathbf{C} = \beta^{-1}\mathbf{I} + \mathbf{D}\, \mathbf{A}^{-1}\, \mathbf{D}^T$ and $\mathbf{I}$ is the identity matrix. To maximize (2.26), an expectation-maximization algorithm is used which provides the point estimates

for the $\boldsymbol{\alpha}$ and β as

$$\alpha_i = \frac{\gamma_i}{\mu_i^2}, \quad i = 0, \ldots, N-1, \tag{2.27}$$

where μ_i is the i-th posterior mean in Eq. (2.24). Here, γ_i is defined as $\gamma_i = 1 - \alpha_i \Sigma_{ii}$ with Σ_{ii} as the i-th diagonal element of the posterior covariance in Eq. (2.25). The noise variance β is estimated as

$$\beta = \frac{\|\mathbf{y} - \mathbf{D}\,\mathbf{x}\|_2^2}{K - \sum_i \gamma_i}. \tag{2.28}$$

Note that $\boldsymbol{\alpha}$ and β are a function of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, while $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are a function of $\boldsymbol{\alpha}$ and β. This therefore suggests an algorithm that iterates between Eqs. (2.24)-(2.28). Once a convergence criterion has been satisfied, the estimated vector $\hat{\mathbf{x}}$ is equal to the posterior mean $\boldsymbol{\mu}$.

This iterative expectation-maximization algorithm produces a highly accurate sparse estimation for sparse vector $\hat{\mathbf{x}}$, but it requires an $O(N^3)$ operation to perform an inversion of a matrix of size $N \times N$ in Eq. (2.25). A fast maximization algorithm has been developed to overcome this problem by analyzing the properties of the marginal likelihood function [61, 63]. For monotonically maximizing the marginal likelihood, this fast algorithm allows an efficient sequential addition and deletion of candidate basis functions (columns of $\mathbf{D}$) until all S relevant basis functions are included. The operational complexity reduces from $O(N^3)$ to $O(N S^2)$, thereby being much more efficient than the original expectation-maximization algorithm, specifically when the signal is truly sparse ($S \ll N$).

2.5 Chapter summary

In this chapter, we have reviewed the CS framework, its underlying premises and principles including sparsity, compressibility, RIP, and coherence. We have also examined several numerical sparse approximation algorithms for signal recovery in noiseless and noisy sensing operations. Furthermore, a Bayesian approach for CS, namely Bayesian compressed sensing is presented in this chapter.

Compressed sensing has increasingly become a powerful framework for several applications, specifically where the signal has a sparse representation in some bases. In the rest of the thesis, we will focus on the applications of CS, Bayesian CS, and sparse representation to TWRI. The next chapter introduces high-resolution TWRI and reviews major CS techniques for TWRI.

Through-the-wall radar imaging and CS-based TWRI techniques

Chapter contents

In the recent area of urban sensing, there is a high demand for high-resolution imaging of stationary targets. Traditionally, this requires the use of wideband or ultra-wideband signals and large array antennas. Employing the ultra-wideband signals and huge array antennas, however, leads to prolonged data acquisition and high cost of operation and processing. Furthermore, for real urban sensing applications, the radar system is typically mounted on a small platform or carried by a human that requires the array antenna to be relatively small and mechanically simple. Thus, imaging techniques that can produce high-quality images

using a reduced measurement set need to be investigated. With the capability of efficiently sensing and reconstructing sparse signals from compressed measurements, compressed sensing is considered as a promising technique for enhancing TWRI.

In this chapter, we first introduce high-resolution through-the-wall radar imaging and describe a backprojection technique for through-the-wall radar image formation. Then, we review major CS applications to TWRI including CS-based image formation and CS-based signal reconstruction. Finally, we discuss the existing limitations and research gaps that are addressed in this research project.

3.1 High-resolution TWRI

High-resolution radar imaging is essential to localize, identify, and track behind wall objects. Two equal-strength scatterers are considered to be resolved if they result in two distinguishably identifiable targets in the scene image [64]. To resolve targets in the formed image, the resolution of a TWRI system must be at least equal to the Rayleigh resolution applied in the downrange and crossrange. The downrange and crossrange resolutions are influenced by the bandwidth of the transmitted signals and the length of the antenna array. More specifically, the downrange resolution δ_d is defined as [2, 65]

$$\delta_d = \frac{c}{2B},\tag{3.1}$$

where c is the speed of light in the air, and B is the bandwidth of the transmitted signal. The crossrange resolution δ_c can be expressed as

$$\delta_c = \frac{r\lambda}{2L},\tag{3.2}$$

where r is the downrange, λ is the signal wavelength, and L is the array length.

For high-resolution TWRI, wideband or ultra-wideband signals are transmitted by either employing a wideband pulse formation or a stepped-frequency

waveform. The former method sends a wideband pulse and uses matched filtering to obtain time-delay information for all reflecting targets in the range. This scheme results in a high data acquisition speed since one measurement covers the whole frequency band, but wideband pulse formation and processing hardware is costly and not flexible. Thus, many radar systems use a stepped-frequency approach to approximate the wideband pulse. In this approach, instead of transmitting a single wideband signal, several mono-frequency signals are transmitted covering the desired frequency band. For each frequency, only the amplitude and phase of the received signal are recorded, by using a comparably inexpensive network analyzer. Stepped-frequency TWRI systems therefore involve transmitting and receiving multiple data samples or multiple stepped-frequencies at each antenna location [1, 66, 67, 68].

The major determinant of the radar system's cross-range resolution is the array antenna. For high-resolution imaging, the larger array antenna needs to be designed using physical or synthetic array aperture. Physical arrays are arrangements of a large number of transceivers covering the desired aperture and thus form big antennas. As a result, all measurements can be acquired simultaneously, reducing the acquisition time. However, a large number of physical transceivers causes the antenna to be large, bulky and costly. Hence, synthetic aperture technique is frequently used to realize a large antenna by moving a single sensor to several positions related to the scene being imaged. Due to the efficiency of designs and deployment, stepped-frequency synthetic aperture radar (SAR) technique is employed in numerous sensing applications. In this research, we investigate the TWRI problems using the stepped-frequency SAR system.

3.2 TWR image formation

In this section, we present a conventional image formation algorithm for producing an image of the scene using a stepped-frequency monostatic TWR system. This radar system illuminates a behind-wall scene by transmitting the stepped-

frequency signals and receiving the radar returns with a transceiver from an initial location. The transceiver antenna is then moved to the next location. The same operation is repeated by moving the transceiver horizontally parallel to the front wall at a standoff distance z_{off}. Assume that a linear antenna aperture comprising M elements has been realized. At each antenna location, the transceiver transmits a stepped-frequency signal consisting of N frequencies equally spaced in the frequency bandwidth $f_{N-1} - f_0$,

$$f_n = f_1 + (n-1)\Delta f, \text{ for } n = 1, 2, \ldots, N, \tag{3.3}$$

where f_1 is the lowest frequency, and $\Delta f = (f_N - f_1)/(N-1)$ is the frequency step size.

Suppose that there are P targets in the scene. Let $z_{m,n}$ denote the signal of frequency f_n received at the m-th antenna location, for $m = 0, 1, \ldots, M-1$. The signal $z_{m,n}$ due to the target reflections can be expressed as

$$z_{m,n} = \sum_{p=0}^{P-1} \sigma_p \, \exp(-j2\pi \, f_n \, \tau_{m,p}). \tag{3.4}$$

It is important to note that Eq. (3.4) models for the radar returns due to the targets present in the scene only. Here, σ_p is the reflectivity of the p-th target, and $\tau_{m,p}$ is the round-trip time of the signal from the m-th antenna to the the p-th target. The round-trip travel time $\tau_{m,p}$ is calculated as [69]

$$\tau_{m,p} = \frac{2}{c}(l_{m,p,air1} + \sqrt{\epsilon}\, l_{m,p,wall} + l_{m,p,air2}), \tag{3.5}$$

where ϵ is the dielectric constant of the wall material, $l_{m,p,air1}$, $l_{m,p,wall}$, and $l_{m,p,air2}$ are, respectively, the distances traveled by the signal before, through, and behind the wall from the m-th antenna, see Fig. 3.1.

For image formation, the target space behind the wall is partitioned into a rect-
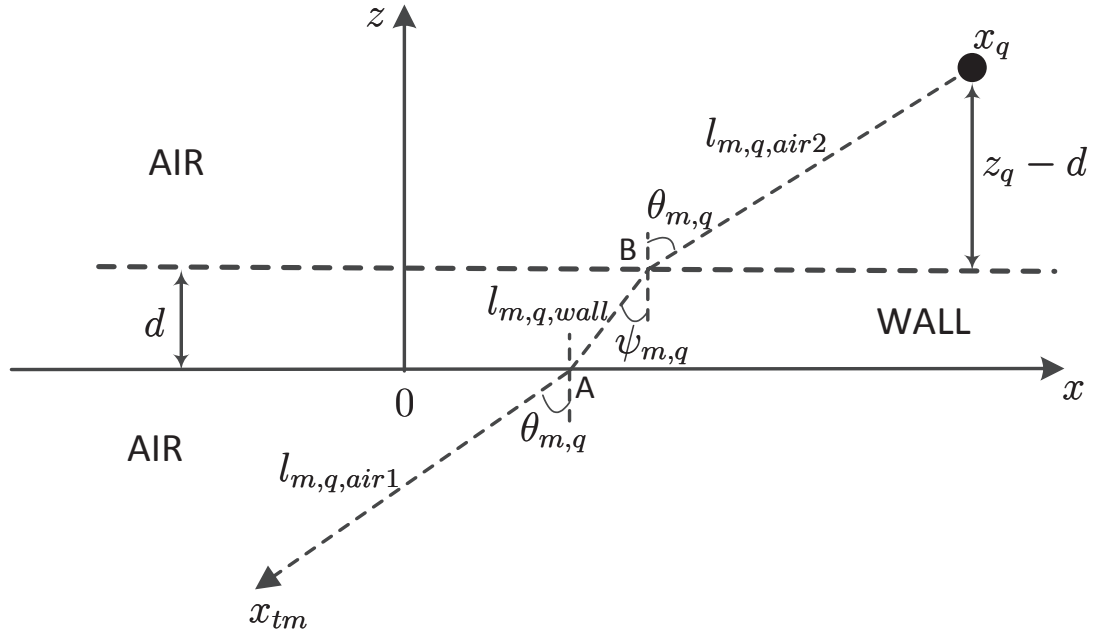
Figure 3.1: Geometry for computation of the focusing delay between the transceiver and the target. Adapted from [70].

angular grid along the crossrange and the downrange. After all frequency measurements from all antennas are acquired, the delay-and-sum (DS) beamforming can be applied to form a complex-valued image by aggregating the measurements $z_{m,n}$ [1, 71]. Let Q be the total number of pixels in the image. The value of the q-th pixel I_q is calculated as

$$I_q = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} z_{m,n} \exp(j2\pi f_n \tau_{m,q}), \tag{3.6}$$

for $q = 0, 1, \ldots, Q-1$. Here $\tau_{m,q}$ is the focusing delay applied to the output of the m-th transceiver for the q-th pixel. The focusing delay is given by replacing the target subscript p with the image pixel subscript q in Eq. (3.5).

In the signal modeling and image formation, it is important to estimate the traveling distances between the transceivers and the pixel locations and subsequently compute the focusing delays. Based on the Snell law, the traveling distances and focusing delays can be calculated. Consider the signal traveling from the m-th antenna transceiver at x_{tm} to the q-th pixel at x_q. This signal goes

through the air-wall-air interfaces and has refraction as shown in Fig. 3.1. By Snell law, the angle of refraction $\psi_{m,q}$ is calculated as

$$\psi_{m,q} = \sin^{-1}(\frac{\sin\theta_{m,q}}{\sqrt{\epsilon}}), \tag{3.7}$$

where $\theta_{m,q}$ is the angle of incident. From the geometry in Fig. 3.1, the distances traveled by the signal before, through, and behind the wall from the m-th antenna are given by

$$l_{m,q,\text{air1}} = \frac{z_{\text{off}}}{\cos\theta_{m,q}},$$
$$l_{m,q,\text{wall}} = \frac{d}{\cos\psi_{m,q}},$$
$$l_{m,q,\text{air2}} = \frac{z_q - d}{\cos\theta_{m,q}}. \tag{3.8}$$

The coordinates of point A in Fig. 3.1 are $(x_{tm} + z_{\text{off}}\tan\theta_{m,q}, 0)$. Applying the cosine law to the triangle with vertices (A,B,x_q), we obtain

$$[x_q - (x_{tm} + z_{off}\tan\theta_{m,q})]^2 = l_{m,q,\text{wall}}^2 + l_{m,q,\text{air2}}^2 - 2l_{m,q,\text{wall}}l_{m,q,\text{air2}}\cos(\pi + \psi_{m,q} - \theta_{m,q}). \tag{3.9}$$

Note that Eqs. (3.7) and (3.9) are transcendental equations and thus can be solved numerically using several methods, such as the Newton-Raphson [1]. After the angle of incident $\theta_{m,q}$ and angle of refraction $\psi_{m,q}$ are obtained, the traveling distances are given using Eq. (3.8), and subsequently the focusing delay $\tau_{m,q}$ is computed as in Eq. (3.5), but replacing the subscript p with q.

It is noted that the conventional DS beamforming is effective for image-based indoor target detection and localization, when a large aperture array and a large signal bandwidth are used [5, 6]. However, its limitation is that the full data volume is required to form a high-quality scene image; otherwise, the image quality deteriorates rapidly with a reduction of measurements. In practical applications, it is vital to generate a scene image using reduced measurements since several

data samples in space and frequency domains can be difficult, or even impossible to obtain. Hence, alternative imaging techniques based on compressed sensing need to be investigated for TWRI.

3.3 TWR image formation using CS

As an innovative and revolutionary idea, CS has become a signal processing tool that has found broad applications including compressive imaging [72, 73, 74], medical imaging [75, 76, 77], source localization [78, 79], and analog-to-information conversion [80, 81]. In radar imaging, CS has been applied to improve sensing systems in several operations [82, 83, 84, 85, 86, 87, 88]. In TWRI, the main objective is to design fast acquisition, low-cost, portable TWRI systems that can retain performances and desirable features as systems operating with full data volume. Over the past decade, CS has been used for TWRI to save data acquisition, reduce computation cost, and improve image formation and fusion [11, 68, 89, 90, 91, 92, 93, 94]. In these methods, CS finds its applications mainly for TWR image formation and data signal reconstruction. In the next subsections, we review these two major CS techniques for TWRI.

3.3.1 CS for image formation

We present a CS-based TWR image formation approach using only a reduced measurement set. The assumptions here are that wall returns have been removed and the scene does not contain large targets or extended objects. Furthermore, the imaging scene contains only P targets, which is far fewer the total number of pixels Q in the entire scene. In other words, the scene can be considered to be sparse, and CS is therefore applicable for reconstructing the sparse image of the scene.

To apply CS, we need to represent the signal model in Eq. (3.4) in a matrix-

vector form. Let $s(q)$ denote an indicator function defined as

$$s(q) = \begin{cases} \sigma_p, & \text{if the } p\text{-th target occupies the } q\text{-th pixel;} \\ 0, & \text{otherwise.} \end{cases} \tag{3.10}$$

Let $\mathbf{z}_m$ denote the column vector containing the frequency measurements collected by the m-th antenna, see Eq. (3.4). Similarly, we denote by $\mathbf{s}$ the lexicographically ordered column vector containing the pixel values of the scene. Let $\mathbf{\Psi}_m$ be an $N \times Q$ matrix whose nq-th element defined as $[\mathbf{\Psi}_m]_{nq} = \exp(-j2\pi f_n \tau_{m,q})$. Here, $\tau_{m,q}$ is the propagation delay between the m-th antenna and the q-th pixel. It follows from Eqs. (3.4) and (3.10) that

$$\mathbf{z}_m = \mathbf{\Psi}_m \, \mathbf{s}. \tag{3.11}$$

By concatenating the received signals at all M antennas, we can write

$$\mathbf{z} = \mathbf{\Psi} \, \mathbf{s}, \tag{3.12}$$

where $\mathbf{z} = [\mathbf{z}_0^T, \ldots, \mathbf{z}_{M-1}^T]^T$ and $\mathbf{\Psi} = [\mathbf{\Psi}_0^T, \ldots, \mathbf{\Psi}_{M-1}^T]^T$.

From Eq. (3.12), if the full data measurements $\mathbf{z}$ are available, the image of the scene $\mathbf{s}$ can be recovered by applying delay-and-sum (DS) beamforming or back-projection [1, 2]. However, in compressed TWR sensing operations, only a partial data set is acquired using a reduced set of M_a antenna locations ($M_a < M$) and a subset of N_f frequencies ($N_f \ll N$). Applying the DS beamforming on the reduced data will result in a degraded image of the scene [89, 95]. Therefore, alternative imaging approaches are needed for forming a high-quality scene image.

Assuming that the number of target pixels are much smaller than the number of total pixels in the scene ($P \ll Q$), from Eq. (3.10), the entries of scene vector $\mathbf{s}$ are nonzero only at the locations occupied by the targets. Hence $\mathbf{s}$ is a P-sparse vector and thus according to CS, a high-quality image can still be generated, although some of the data samples are missing. Let $\mathbf{y}$ denote a vector consisting

of measurements obtained by randomly selecting a subset of M_a antenna positions and N_f frequencies. Mathematically, this can be represented using a projection matrix $\boldsymbol{\Phi}$ with $M_a \times N_f$ rows and $M \times N$ columns. Each row of $\boldsymbol{\Phi}$ has exactly one entry with a value of 1 at a position determined by the selected antennas and frequency bins; all other entries of the row are equal to 0. Thus, the reduced measurement vector $\mathbf{y}$ is expressed as

$$\mathbf{y} = \boldsymbol{\Phi}\, \mathbf{z} = \boldsymbol{\Phi}\, \boldsymbol{\Psi}\, \mathbf{s} = \mathbf{D}\, \mathbf{s}, \tag{3.13}$$

where $\mathbf{D} = \boldsymbol{\Phi}\, \boldsymbol{\Psi}$.

From compressed vector $\mathbf{y}$, the scene $\mathbf{s}$ is reconstructed by solving the following ℓ_1-norm minimization:

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \|\mathbf{s}\|_1 \quad \text{subject to} \quad \mathbf{y} = \mathbf{D}\, \mathbf{s}. \tag{3.14}$$

The problem (3.14) can be recast into a quadratically-constrained ℓ_1 minimization as

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \|\mathbf{s}\|_1 \quad \text{subject to} \quad \|\mathbf{y} - \mathbf{D}\, \mathbf{s}\|_2 \le \epsilon, \tag{3.15}$$

where ϵ is a bound noise. This parameter can be estimated using a cross-validation technique [47, 48]. The problem (3.15) is known as basis pursuit denoising (BPDN), and hence several BPDN solvers [15, 38, 44], NESTA [45], or SPAMS [46] can be used to obtain the image $\hat{\mathbf{s}}$. Greedy algorithms, such as orthogonal matching pursuit (OMP) [49, 50] can also be applied but they often require more measurements compared with the standard convex optimization methods.

This CS-based TWR image formation method was first presented in [89, 95]. It has been reported that the CS-based approach can yield a high-resolution image of the scene provided that at least $O(P \log(Q/P))$ measurements are available. Note that this CS-based approach assumes the strong wall returns have been completely removed prior to applying CS, otherwise the sparsity assumption of

scene **s** is violated and the CS application is ineffective or even impossible.

3.3.2 CS for data recovery

Here, we review the application of CS for recovering the missing data samples at each antenna. Regular image formation methods, such as DS beamforming, can then be implemented based on the recovered data samples, as if they were physically acquired. In other words, CS is used for generating the same image as one obtained by backprojection, but with fewer data measurements.

Let $\mathbf{z}_m$ denote the column vector containing the frequency measurements collected by the m-th antenna, see Eq. (3.4). Let $\mathbf{u}_m$ be the discrete range profile. Suppose that the range of interest is partitioned into L equally spaced range cells. The range vector $\mathbf{u}_m$ is an L-dimensional column vector defined as

$$[\mathbf{u}_m]_l = \begin{cases} \sigma_p, & \text{if } \tau_l = \tau_{m,p}, \\ 0, & \text{otherwise.} \end{cases} \quad (3.16)$$

Here, τ_l, and $\tau_{m,p}$ are, respectively, the two-way signal traveling time from the m-th antenna location to the l-th range cell and to the p-th target. Then, the measurement signal $\mathbf{z}_m$ is related to the range $\mathbf{u}_m$ as

$$\mathbf{z}_m = \mathbf{\Psi}\, \mathbf{u}_m, \quad (3.17)$$

where $\mathbf{\Psi}$ is an $N \times L$ matrix defined as

$$[\mathbf{\Psi}]_{nl} = \exp(-j2\pi\, f_n\, \tau_l). \quad (3.18)$$

In indoor imaging, the number of targets is significantly small compared to the entire discrete spatial positions of the scene. Therefore, the range vectors $\mathbf{u}_m$ are sparse, and CS theory can be applied to reconstruct the range profiles from compressive measurements. Let $\mathbf{y}_m$ be a K-dimensional vector consisting of data

samples randomly chosen from $\mathbf{z}_m$. Vector $\mathbf{y}_m$ is mathematically expressed as

$$\mathbf{y}_m = \mathbf{\Phi}\,\mathbf{z}_m, \tag{3.19}$$

where $\mathbf{\Phi}$ is a K-by-N random selection matrix containing only one non-zero element (a value of 1) in each row, and $K \ll N$. From Eq. (3.17), it follows that

$$\mathbf{y}_m = \mathbf{\Phi}\,\mathbf{\Psi}\,\mathbf{u}_m = \mathbf{D}\,\mathbf{u}_m, \tag{3.20}$$

where $\mathbf{D} = \mathbf{\Phi}\,\mathbf{\Psi}$ is a $K \times L$ dictionary matrix. Let ϵ be a noise bound. The sparse vectors $\mathbf{u}_m$ ($m = 0, 1, ..., M-1$) are recovered by solving the following CS inverse problems:

$$\hat{\mathbf{u}}_m = \arg\min_{\mathbf{u}_m} \|\mathbf{u}_m\|_1 \text{ subject to } \left\|\mathbf{D}\,\mathbf{u}_m - \mathbf{y}_m\right\|_2 \leq \epsilon. \tag{3.21}$$

The above M inverse problems can be solved separately to recover M range profiles. This approach has been proposed in previous work [68, 91]. Alternatively, the M compressed measurement vectors can be concatenated into a single composite measurement vector, which is then used to recover all the range profiles simultaneously. This technique, however, increases the complexity of the problem by M-fold. Once all the sparse vectors $\hat{\mathbf{u}}_m$, for $m = 0, 1, \ldots, M-1$ have been obtained, the recovered signals $\hat{\mathbf{z}}_m$ are estimated using Eq. (3.17): $\hat{\mathbf{z}}_m = \mathbf{\Psi}\,\hat{\mathbf{u}}_m$.

3.4 Discussions

Compressed sensing has been applied and shown to be a promising signal processing framework for TWRI, but several challenges need to be addressed for enhancing CS applications to urban sensing. One of the challenge is that in certain applications, the scene contains large or extended targets, such as human and furniture. This makes the point-like target assumption invalid and reduces the sparsity of the scene, thereby impeding the direct CS application. To overcome this problem, we propose a two-stage CS-based approach, presented in Chapter 4,

for TWR image formation. It incorporates a dictionary used to sparsely represent the scene. Experimental results and analysis demonstrate that incorporating the sparsifying dictionary makes the CS technique to be more effective even when the measurement set is drastically reduced.
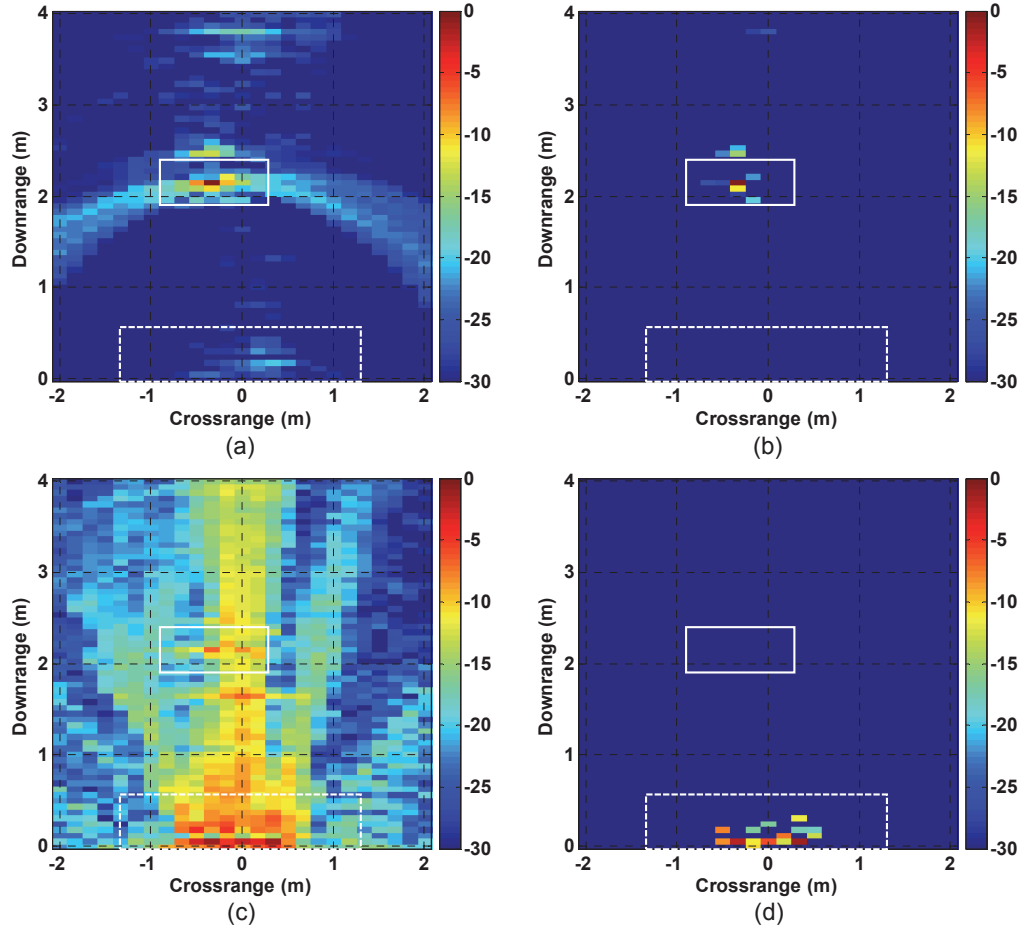
Figure 3.2: Images of a dihedral scene formed by: (a)–(b) DS beamforming and CS-based techniques after wall-clutter mitigation; (c)–(d) DS beamforming and CS-based techniques without wall clutter mitigation. Wall and target regions are indicated with dashed and solid rectangles, respectively.

Imaging behind-the-wall targets in the presence of strong wall clutter is a challenging but vital task in compressed TWR sensing. Most CS-based TWRI approaches assume that the wall radar returns have been removed completely from the received signals. As a result, the signal-to-clutter ratio is enhanced and targets can be localized in the formed images. For illustration, Figs. 3.2(a)–(b) show the images of a dihedral scene formed by the DS beamforming and CS-

based techniques after wall clutter mitigation. It is observed that the dihedral target is obviously detected with the strong intensity. However, in practical applications, the wall reflections typically dominate the target signals and reduce the sparsity of the signal significantly. Consequently, the direct CS application is not effective. As an example, Figs. 3.2(c)–(d) present the dihedral scene images formed by the DS beamforming and CS-based approaches without wall clutter mitigation. Clearly, the wall clutter obscures the target, rendering target detection difficult or even impossible. Thus, it is crucial to investigate new techniques that enable scene reconstruction while mitigate wall clutter and reverberations. In this research, we develop a Bayesian scene reconstruction approach, presented in Chapter 5, which jointly performs wall clutter suppression and image formation in the CS context.

Multi-view and multi-polarization TWRI techniques have recently shown to improve target detection and localization by combining multiple data sets collected from multiple vantage points or channels. Collecting data at several vantage points improves imaging visibility, but also leads to prolonged data acquisition time and expensive hardware. Since most existing CS techniques are proposed for single-view or single-polarization model, deviations based on CS need to be investigated for multi-polarization and multi-view TWRI. In this project, we address the problem of multi-channel TWRI using joint Bayesian CS framework. The proposed Bayesian approach, presented in Chapter 6, aims to relax constraints on signal sampling schemes and logistic difficulties in data acquisition. It considers the wall clutter mitigation and takes into account the inter-channel dependences for improving target identification and alleviating background clutter.

3.5 Chapter summary

In this chapter, we have reviewed high-resolution through-the-wall radar imaging, TWR image formation, and CS techniques for TWRI. It has been shown that generating high-resolution TWRI is vital for target detection and localization, but

conventional TWRI methods are not flexible for urban sensing applications due to system constraints. Compressed sensing techniques have shown to be a promising signal processing framework for enhancing TWRI, specifically under sensing operations where the full data volume cannot be attained. Assuming the scene contains point-like targets and the signals consist of target returns only, CS can be applied to reconstruct a high-quality image of the scene. Moreover, CS can be employed for data recovery, followed by conventional beamforming methods for image formation. We also discuss the weaknesses and challenges of the existing CS-based techniques. The remaining chapters of this thesis present several new imaging approaches based on CS and Bayesian CS for handling the problems and challenges in urban sensing operations.

Chapter 4

Two-stage through-the-wall radar image formation using compressive sensing

Chapter contents

Abstract

This chapter introduces a robust image formation approach for through-the-wall radar imaging. The proposed approach consists of two stages involving compres-

sive sensing followed by delay-and-sum (DS) beamforming. In the first stage, CS is used to reconstruct a complete set of measurements from a small subset collected with a reduced number of transceivers and frequencies. DS beamforming is then applied to form the image using the reconstructed measurements. To promote sparsity of the CS solution, an overcomplete Gabor dictionary is employed to sparsely represent the imaged scene. The new approach requires far fewer measurement samples than the conventional delay-and-sum beamforming and CS-based TWRI methods to reconstruct a high-quality image of the scene. Experimental results based on simulated and real data demonstrate the effectiveness and robustness of the proposed two-stage image formation technique, especially when the measurement set is drastically reduced.

4.1 Introduction

Through-the-wall radar imaging is an emerging technology with considerable research interest and important applications in surveillance and reconnaissance for both civil and military missions [1, 2, 3, 69, 71, 96]. To deliver high-resolution radar images in both range and crossrange, TWRI systems use wideband signals and large aperture arrays (physical or synthetic). This leads to prolonged data acquisition and high computational complexity because a large number of samples need to be processed. New approaches for TWRI are therefore needed to obtain high-quality images from fewer data samples at a faster speed. To this end, this chapter proposes a new approach using compressive sensing for through-the-wall radar imaging. Compressive sensing is used here to reconstruct a full measurement set, which is then employed for image formation using delay-and-sum beamforming.

Compressive sensing enables a sparse signal to be reconstructed using considerably fewer data samples than what is required by the Nyquist-Shannon theorem [14, 15, 17]. In through-the-wall radar imaging, the objective of ap-

---

The content of this chapter has been published as V. H. Tang, A. Bouzerdoum, and S. L. Phung, "Two-stage through-the-wall radar image formation using compressive sensing," *SPIE Journal of Electronic Imaging*, vol. 22, no. 2, pp. 021 006–1–021 006–10, 2013.

plying CS is to speed up data acquisition and achieve high-resolution imaging [89, 90, 91, 92, 93, 95]. So far, the application of CS in TWRI can be divided into two main categories. In the first category, CS is applied to reconstruct the imaged scene directly by solving an ℓ_1 optimization problem or using a greedy reconstruction algorithm [89, 90, 92, 93, 95]. In the second category, CS is employed in conjunction with the traditional beamforming methods. In other words, CS is applied to reconstruct the full data volume, and the conventional imaging methods, such as DS beamforming, are then used to form the image of the scene [91]. By exploiting CS, the latter approach enables conventional beamforming methods to reconstruct high-quality images from reduced data samples. Moreover, adopting a conventional image formation approach produces images suitable for target detection and classification tasks, which typically follow the image formation step.

In [91], the full measurement set is recovered from the range profiles obtained by solving a separate CS problem at each sensor location. CS is applied in the temporal frequency domain only, leaving uncompressed sensing in the spatial domain. To recover the full measurement set, several CS problems are solved independently—one for each sensing location. There are also limitations in reducing the measurements along the temporal frequencies. Since the target radar-cross-section depends highly on signal frequency, significant reduction in transmitted frequencies will lead to deficient information about the target [90]. Thus, to guarantee accurate reconstruction, imaging the scene with extended targets may require an increase in the number of measurements [2, 97].

The conventional beamforming methods have been shown to be very effective for image-based indoor target detection and localization when using a large aperture array and large signal bandwidth [4, 5, 70, 98, 99, 100, 101, 102]. However, a limitation of the traditional beamforming methods is that they require the full data volume to form a high-quality image; otherwise, the image quality deteriorates rapidly with a reduction of measurements. The question is then how to exploit

the advantages of the traditional beamforming methods to obtain high-quality images from a reduced set of measurements.

To answer the aforementioned question and address the limitation of existing CS-based imaging methods, this chapter proposes a new CS approach for TWRI, whereby a significant reduction in measurements is achieved by compressing both the transmitted frequencies and the sensor locations. First, CS is employed to restore the full measurement set. Then DS beamforming is applied to reconstruct the image of the scene. To increase sparsity of the CS solution, an overcomplete Gabor dictionary is used for sparse representation of the imaged scene; Gabor dictionaries have been shown to be effective for image sparse decomposition and representation [103, 104, 105]. In the proposed approach, fast data acquisition is achieved by reducing both the number of transceivers and transmitted frequencies used to collect the measurement samples. In [91], data collection was performed at all antenna locations, using a reduced set of frequencies only. In contrast, the proposed approach achieves further measurement reduction by subsampling both the number of frequencies and antenna locations used for data collection. Furthermore, to satisfy the sparsity assumption, a Gabor dictionary is incorporated in the scene representation. In [92], a wavelet transform was used as a sparsifying basis for the scene. However, our preliminary experiments show that the performance is highly dependent on the particular wavelet function used. We also found that wavelets offer no significant advantage over Gabor basis in the problem of through-the-wall radar image formation. Finally, we should note that there are several approaches that have been proposed for wall clutter mitigation in TWRI [96, 106], including recent successful CS-based techniques [107, 108]. In this chapter, we assume that wall clutter can be removed using any of those techniques, or the background scene is available to perform background subtraction.

The remainder of the chapter is organized as follows. Section 4.2 gives a brief background of compressive sensing theory. Section 4.3 presents TWRI using

delay-and-sum beamforming, and describes the proposed approach for TWRI image formation. Section 4.4 presents experimental results and analysis. Section 4.5 gives concluding remarks.

4.2 CS background

Compressive sensing is an innovative and revolutionary idea that offers joint sensing and compression for sparse signals [14, 15, 16, 17]. Consider a P-dimensional signal $\mathbf{x}$ to be represented using a dictionary $\mathbf{\Psi} \in \mathbb{R}^{P \times Q}$ with Q atoms. The dictionary is assumed to be overcomplete, that is, $Q > P$. Signal $\mathbf{x}$ is said to be K-sparse if it can be expressed as

$$\mathbf{x} = \mathbf{\Psi}\,\boldsymbol{\alpha}, \tag{4.1}$$

where $\boldsymbol{\alpha}$ is a column vector with K nonzero components, i.e., $K = \|\boldsymbol{\alpha}\|_0$. Stable reconstruction of a sparse $\boldsymbol{\alpha}$ requires K to be significantly smaller than P.

Using a projection matrix $\mathbf{\Phi}$ of size $L \times P$, where $K < L \ll P$, we can obtain an L-dimensional measurement vector $\mathbf{y}$ as follows:

$$\mathbf{y} = \mathbf{\Phi}\,\mathbf{x}. \tag{4.2}$$

The original signal $\mathbf{x}$ can be reconstructed from $\mathbf{y}$ by exploiting its sparsity. Among all $\boldsymbol{\alpha}$ satisfying $\mathbf{y} = \mathbf{\Phi}\,\mathbf{\Psi}\,\boldsymbol{\alpha}$, we seek the sparsest vector, and then obtain $\mathbf{x}$ using (4.1). This signal reconstruction requires solving the following problem:

$$\min \|\boldsymbol{\alpha}\|_0 \quad \text{subject to} \quad \mathbf{y} = \mathbf{\Phi}\,\mathbf{\Psi}\,\boldsymbol{\alpha}. \tag{4.3}$$

Problem (4.3) is known to be NP-hard [109]. Alternatively, the problem can be cast into an ℓ_1 regularization problem:

$$\min \|\boldsymbol{\alpha}\|_1 \quad \text{subject to} \quad \left\|\mathbf{y} - \mathbf{\Phi}\,\mathbf{\Psi}\,\boldsymbol{\alpha}\right\|_2 \leq \epsilon, \tag{4.4}$$

where ϵ is a small constant. Several optimization methods, including ℓ_1-optimization

[45], basis pursuit [110], and orthogonal matching pursuit [49], have been proposed that produce stable and accurate solutions.

4.3 Two-stage TWRI using CS

In this section, we introduce the proposed two-stage TWRI approach based on compressive sensing. The main steps of the proposed approach are as follows. First, compressive measurements are acquired using a fast data acquisition scheme that requires only a reduced set of antenna locations and frequency bins. An additional Gabor dictionary is incorporated into the CS model to sparsely represent the scene. Next, the full TWRI data samples are recovered, and then conventional DS technique is applied to generate the scene image. In this section, we first give a brief review of the conventional delay-and-sum beamforming method for image formation in Section 4.3.1, before presenting the new image formation approach in Section 4.3.2.

4.3.1 TWRI using delay-and-sum beamforming

Consider a stepped-frequency monostatic TWRI system that uses M transceivers and N narrowband signals to image a scene containing R targets. The signal received at the m-th antenna location and n-th frequency, f_n, is given by

$$z_{m,n} = \sum_{r=1}^{R} \sigma_r(f_n) \, \exp\left\{-j2\pi \, f_n \, \tau_{m,r}\right\}, \tag{4.5}$$

where $\sigma_r(f_n)$ is the reflection coefficient of the r-th target for the n-th frequency, and $\tau_{m,r}$ is the round-trip travel time of the signal from the m-th antenna location to the r-th target location. In the stepped-frequency approach, the frequency bins f_n are uniformly distributed over the entire frequency band, with a step size Δf:

$$f_n = f_1 + (n-1)\Delta f, \text{ for } n = 1, 2, \ldots, N, \tag{4.6}$$

where f_1 is the first transmitted frequency.

The target space behind the wall is partitioned into a rectangular grid, with N_x pixels along the crossrange direction and N_y pixels along the downrange direction. Using delay-and-sum beamforming, a complex image is formed by aggregating the measurements $z_{m,n}$. The value of the pixel at location (x, y) is computed as follows:

$$I(x, y) = \frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} z_{m,n} \exp\{j2\pi f_n \tau_{m,(x,y)}\}, \tag{4.7}$$

where $\tau_{m,(x,y)}$ is the focusing delay between the m-th transceiver and the target located at the pixel position (x, y). Assuming that the wall thickness and relative permittivity are known, the focusing delay can be calculated using Snell's law, the distance of the transceiver to the front wall, and the distance of the target to the back wall, see [95, 99], and [111].

4.3.2 Two-stage TWRI

Let $\mathbf{z}$ be the column vector obtained by stacking the data samples $z_{m,n}$ in (4.5), where $m = 1, 2, \ldots, M$ and $n = 1, 2, \ldots, N$. Let s_{xy} be an indicator function defined as

$$s_{xy} = \begin{cases} \sigma_r, & \text{if a target } r \text{ exists at the } xy\text{-th pixel;} \\ 0, & \text{otherwise.} \end{cases} \tag{4.8}$$

The elements s_{xy} are then lexicographically ordered into a column vector $\mathbf{s}$. The magnitude of each element in $\mathbf{s}$ reflects the significance of a point in the scene. From (4.5), the full measurement vector $\mathbf{z}$ can be represented as

$$\mathbf{z} = \boldsymbol{\Psi}\, \mathbf{s}, \tag{4.9}$$

where $\boldsymbol{\Psi}$ is an overcomplete dictionary, which depends on the target scene, the antenna locations, and the transmitted frequencies. More precisely, $\boldsymbol{\Psi}$ is a matrix with $(M \times N)$ rows, and $(N_x \times N_y)$ columns. The entry at row r and column c is given as

$$\Psi_{r,c} = \exp\left\{-j2\pi f_n \tau_{m,(x,y)}\right\}, \tag{4.10}$$
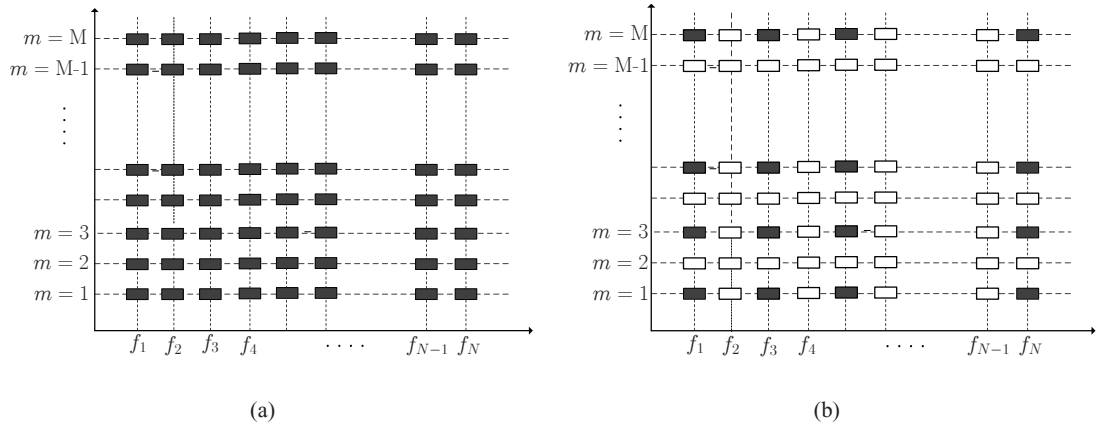
Figure 4.1: Data acquisition for TWRI: (a) Conventional radar imaging scheme; (b) TWRI based on CS. The vertical axis represents the antenna location, and the horizontal axis represents the transmitted frequency. The filled rectangles represent the acquired data samples.

where $r = (m - 1) \times N + n$, and $c = (x - 1) \times N_y + y$.

To reduce the data acquisition time and computational complexity, we propose acquiring only a small number of samples, represented by vector $\mathbf{y}$. The measurements in $\mathbf{y}$ are obtained by selecting only a subset of M_a antenna locations and N_f frequencies. In this chapter, the reduced antenna locations are uniformly selected, and at each selected antenna location, the same number of frequency bins are regularly sub-sampled. This fast data acquisition scheme leads to stable image quality and is more suitable for hardware implementation. Figure 4.1(a) shows the conventional radar imaging where full data samples are acquired. Figure 4.1(b) illustrates the space-frequency sub-sampling pattern used in the proposed approach.

Mathematically, the CS data acquisition can be represented using a projection matrix $\mathbf{\Phi}$ with $(M_a \times N_f)$ rows and $(M \times N)$ columns. Each row of $\mathbf{\Phi}$ has only one non-zero entry with a value of 1 at a position determined by the selected antenna locations and frequency bins. Thus, the reduced measurement vector $\mathbf{y}$ can be expressed as

$$\mathbf{y} = \mathbf{\Phi}\,\mathbf{z} = \mathbf{\Phi}\,\mathbf{\Psi}\,\mathbf{s} = \mathbf{A}\,\mathbf{s}, \tag{4.11}$$

where $\mathbf{A} = \mathbf{\Phi}\,\mathbf{\Psi}$.

44

In practical situations, the sparsity of the scene behind wall is reduced because of multipath propagations, wall reflections and the presence of extended objects, such as people or furniture. Therefore, the sparsity assumption of vector $\mathbf{s}$ may be violated. To overcome this problem, an additional overcomplete dictionary is employed to sparsely represent $\mathbf{s}$. In our approach, a Gabor dictionary is used. Let $\mathbf{W}$ be the synthesis operator for the signal expansion. Thus, the vector $\mathbf{s}$ can be expressed as

$$\mathbf{s} = \mathbf{W}\,\boldsymbol{\alpha}. \tag{4.12}$$

Substituting (4.12) into (4.11) yields

$$\mathbf{y} = \mathbf{A}\,\mathbf{W}\boldsymbol{\alpha}. \tag{4.13}$$

For noisy radar signals, the compressive measurement vector $\mathbf{y}$ is modeled as

$$\mathbf{y} = \mathbf{A}\,\mathbf{W}\,\boldsymbol{\alpha} + \mathbf{v}, \tag{4.14}$$

where $\mathbf{v}$ is the noise component.

The full data volume can be recovered by two techniques: the synthesis method and the analysis method. In the synthesis technique, the problem is cast as follows:

$$\min \|\boldsymbol{\alpha}\|_1 \ \ \text{subject to} \ \ \left\|\mathbf{y} - \mathbf{A}\,\mathbf{W}\,\boldsymbol{\alpha}\right\|_2 \le \epsilon \tag{4.15}$$

Once the coefficient $\boldsymbol{\alpha}$ has been obtained by solving the optimization problem, the full TWRI data samples are obtained, using (4.9) and (4.12),

$$\mathbf{z} = \boldsymbol{\Psi}\,\mathbf{s} = \boldsymbol{\Psi}\,\mathbf{W}\,\boldsymbol{\alpha}. \tag{4.16}$$

In the analysis technique, the problem is formulated as

$$\min \left\|\mathbf{W}^{-1}\,\mathbf{s}\right\|_1 \ \ \text{subject to} \ \ \left\|\mathbf{y} - \mathbf{A}\,\mathbf{s}\right\|_2 \le \epsilon. \tag{4.17}$$

By solving this optimization problem, we obtain the vector **s** directly, which can be used to reconstruct the full measurement vector **z**, see (4.9).

Note that it was suggested in [112] that the analysis technique is less sensitive to noise, compared to the synthesis technique. In our approach, we use the analysis technique for solving the CS problem. After reconstructing the full measurement vector **z**, we apply the conventional delay-and-sum beamforming to generate the scene image as described in Section 4.3.1.

4.4 Experimental results and analysis

In this section, we evaluate the proposed approach using both synthetic and real TWRI data sets. First, the performance of the proposed approach is investigated in Section 4.4.1 using synthetic data. Then, the experimental results on real data are presented in Section 4.4.2, along with the TWRI experimental setup for radar signal acquisition.

4.4.1 Synthetic TWRI data

Data acquisition is simulated for a stepped-frequency radar system, with a frequency range between 0.7 and 3.1 GHz and a frequency step of 12 MHz. Therefore, the number of frequency bins used is $N = 201$. The scene is illuminated with an antenna array of length 1.24 m and an inter-element spacing of 0.022 m, which means the number of transceivers used is $M = 57$. The full data volume **z** comprises $M \times N = 57 \times 201 = 11{,}457$ samples. Our goal is to acquire much fewer data samples without degrading the quality of the image.

The TWRI system is placed in front of a wall at a standoff distance of $Z_{\text{off}} = 1.5$ m. The thickness and relative permittivity of the wall are $d = 0.143$ m and $\epsilon_r = 7.6$, respectively. The downrange and crossrange of the scene extend from 0 m to 6 m, and from −2 m to 2 m, respectively. The pixel size is equal to the Rayleigh resolution of the radar, which gives an image of size 97×65 pixels. In this experiment, three extended targets (each covering 4 pixels) are placed behind the wall at
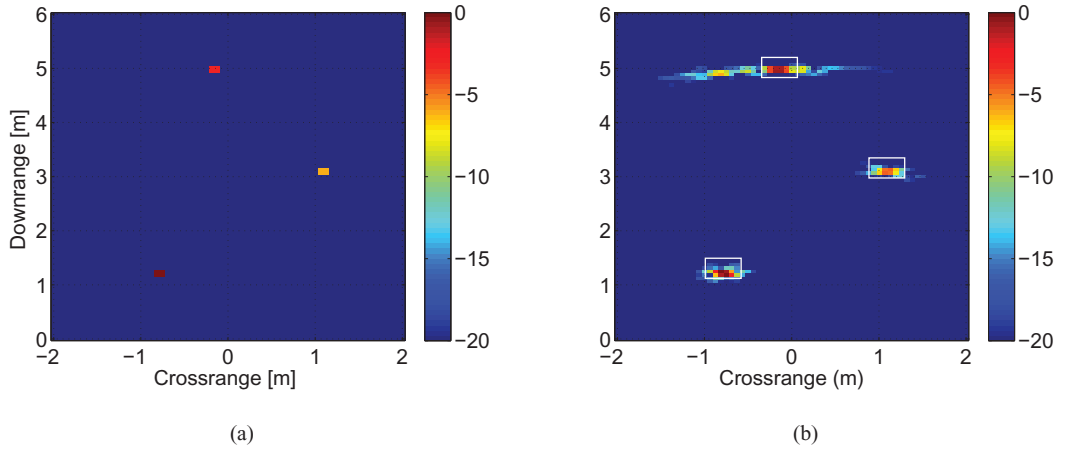
Figure 4.2: The behind-the-wall scene: (a) Ground-truth image; (b) DS image formed using full volume of data samples.

positions $p_1 = (1.21 \text{ m}, -0.78 \text{ m})$, $p_2 = (3.09 \text{ m}, 1.09 \text{ m})$, and $p_3 = (4.96 \text{ m}, -0.16 \text{ m})$. The reflection coefficients are considered to be independent of signal frequency: $\sigma_1 = 1$, $\sigma_2 = 0.5$ and $\sigma_3 = 0.7$, respectively. In our experiment, the first-order method Nesta is used to solve the CS optimization problem with the analysis technique because of its robustness, flexibility, and speed. More details about the Nesta solver can be found in [45]. Here, a dictionary consisting of the complex Gabor functions is used for sparse decomposition of the scene [105].

The peak-signal-to-noise ratio (PSNR) is used to evaluate the quality of the reconstructed images:

$$\text{PSNR} = 20 \log_{10}(I_{\max}/\text{RMSE}), \tag{4.18}$$

where $I_{\max}$ denotes the maximum pixel value, and RMSE is the root-mean-square error between the reconstructed and the ground-truth images. The performance of the proposed approach in the presence of noise is evaluated by adding white Gaussian noise to the received signal.

Figure 4.2 shows the ground-truth image and the DS beamforming image reconstructed using the full measurement volume. Note that in this chapter, all output images are normalized by the maximum image intensity. The true target

position is indicated with a solid white rectangle. Figure 4.3 illustrates the images
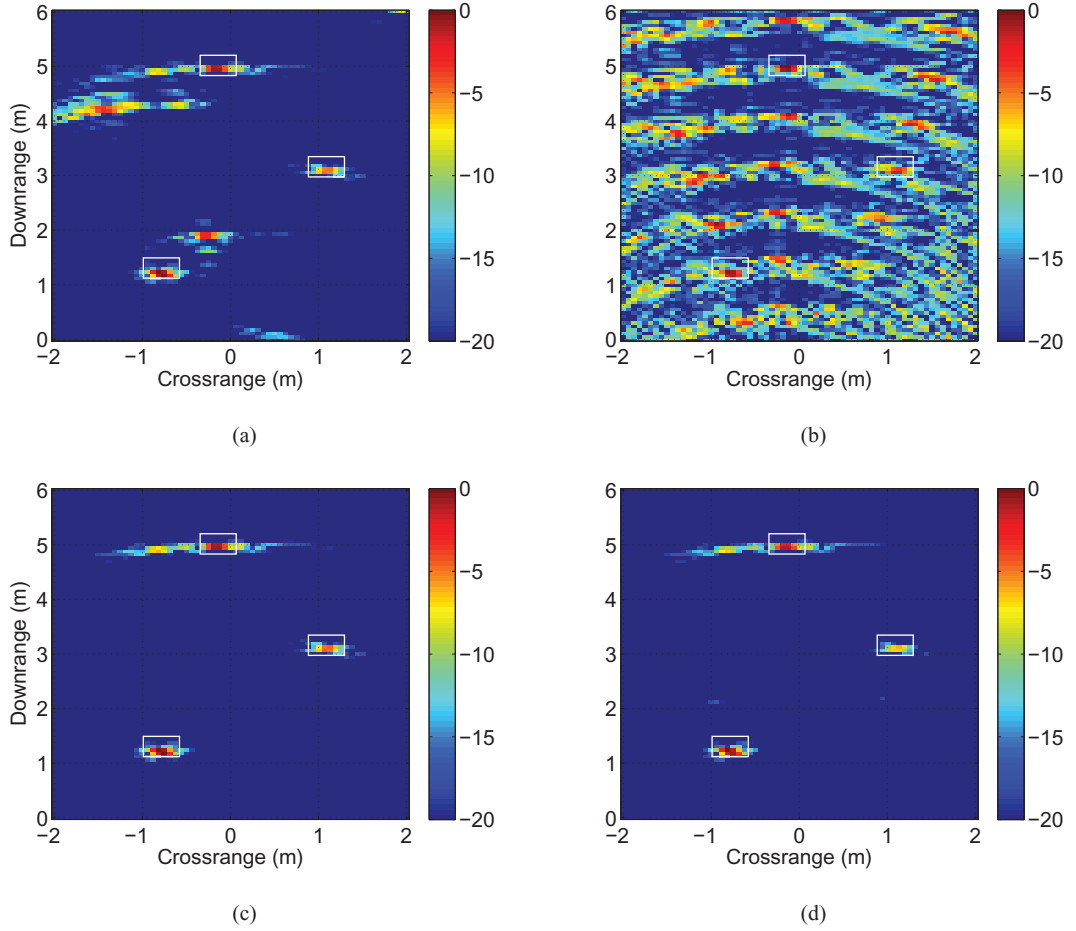
Figure 4.3: Scene images formed by different settings: (a) DS using 12% full data volume; (b) DS using 1% full data volume; (c) proposed approach using 12% full data volume; (d) proposed approach using 1% full data volume. The signal is corrupted by the noise with SNR = 20 dB.

formed with reduced subsets of measurements (12% and 1%), using DS beamforming (Fig. 4.3(a) and 4.3(b)) and the proposed approach (Fig. 4.3(c) and 4.3(d)). Here, the received signals are corrupted by additive white Gaussian (AWG) noise with SNR = 20 dB. Compared to the image obtained using DS beamforming with all measurement samples (Fig. 4.2(b)), the images produced using DS beamforming with reduced data samples (Figs. 4.3(a) and 4.3(b)) deteriorate significantly in quality and contain many false targets. By contrast, Figs. 4.3(c) and 4.3(d) show that images obtained with the proposed approach, using the same reduced datasets, suffer little or no degradation. These results demonstrate that the pro-

posed approach performs significantly better than the standard DS beamforming when the number of measurements is reduced significantly. Furthermore, the images produced by the proposed approach using the reduced data samples have a similar visual quality as the images formed by the standard DS beamforming method using the full data volume.
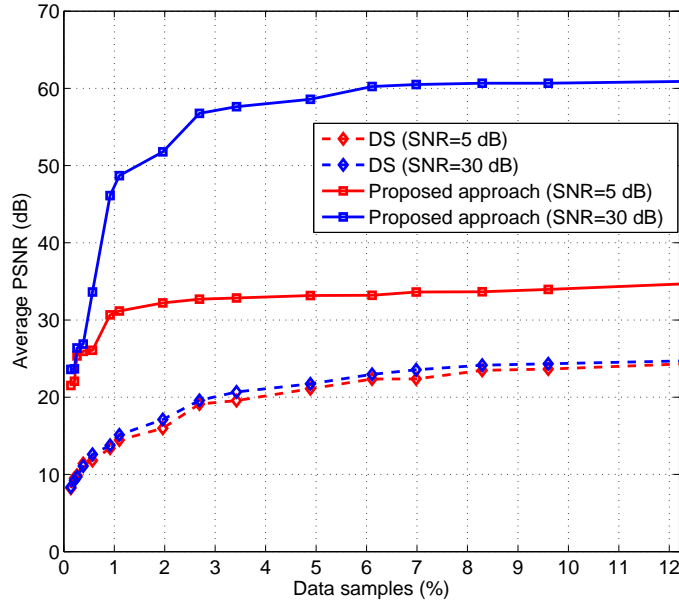
Figure 4.4: The PSNR of images created by the standard DS (dashed lines) and the proposed approach (solid lines).

To evaluate the robustness of the proposed approach in the presence of noise, the measurement signals are corrupted with AWG with SNR equal to 5 dB and 30 dB. Figure 4.4 presents the average PSNR of the reconstructed images as a function of the ratio between the reduced measurement set and the full dataset. The figure clearly shows that the images formed with the proposed approach have considerably higher PSNR than the images formed with the standard DS beam-forming, using the same measurements. This is because the proposed approach reconstructs the full data samples using CS, before applying DS beamforming.

To compare the performance of different imaging methods, we used 3 antenna locations and 40 uniformly selected frequencies, which represents 1% of the total data volume. Figure 4.5 shows the results obtained by different imaging methods.

Figure 4.5(a) shows the CS image reconstructed with the method proposed in [89]. Although the targets can easily be located, there are many false targets in the image. Figure 4.5(b) illustrates the image formed with the method presented in [91]; this image is considerably degraded with the presence of heavy clutter. The reason is that the imaging method in [91] is not able to restore the full data volume from a reduced set of antenna locations. Figure 4.5(c) and 4.5(d) show the images formed with the proposed approach using wavelet and Gabor sparsifying dictionaries, respectively. Here, the wavelet family is the dual-tree complex wavelet transform. It can clearly be observed that the image formed using the Gabor dictionary contains less clutter; however, both dictionaries yield high-quality images even with a significant reduction in the number of collected measurements.

In the next experiment, only the frequency samples are reduced; the data is collected at all antenna locations, using $M = 57$ transceivers. The reduced dataset represents 20% of the full data volume. Figure 4.6 presents the images formed using different approaches: standard CS method [89], the temporal frequency CS method [91], the proposed method with a wavelet dictionary, and the proposed method with a Gabor dictionary. It can be observed from Fig. 4.6(b) that there is a substantial improvement in the performance of the temporal frequency CS method [91]. This is because when using all antenna locations, this imaging method can obtain the full data volume for forming the image. However, the proposed method yields images with less clutter, using both wavelet and Gabor dictionaries.

In summary, the experimental results on synthetic TWRI data demonstrate that the proposed approach produces high-quality images using far fewer measurements by applying CS data acquisition in both frequency domain and spatial domain. It performs better than the conventional DS and CS-based TWRI methods, especially when the number of measurements is drastically reduced.
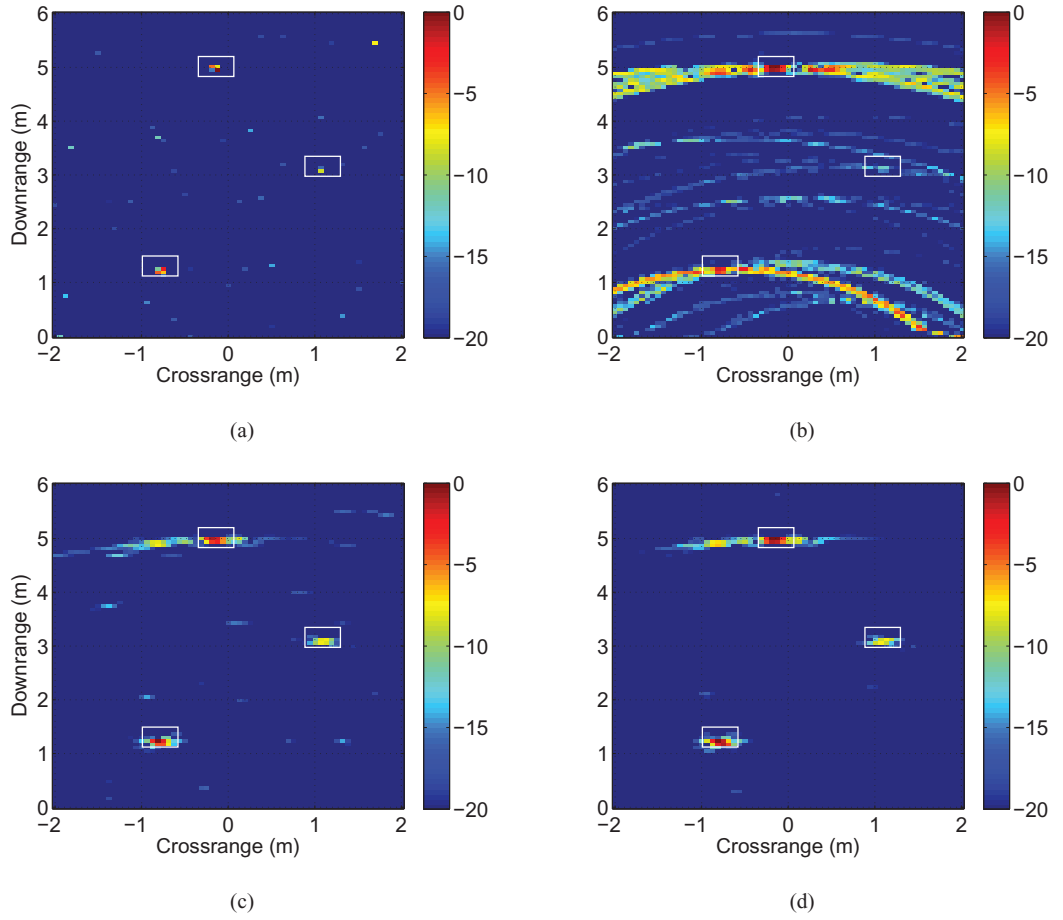
Figure 4.5: Scene images formed by different imaging approaches: (a) CS image formed by method in [89] ; (b) DS image formed by method in [91]; (c) DS image formed by the proposed approach with wavelet dictionary; (d) DS image formed by the proposed approach with Gabor dictionary. The measurements made up 1% of full data volume. The signal is corrupted by the noise with SNR = 10 dB.

4.4.2   Real TWRI data

In this experiment, the proposed approach is evaluated on real TWRI data. The data used in this experiment were collected at the Radar Imaging Laboratory of the Center for Advanced Communications, Villanova University, USA. The radar system was placed in front of a concrete wall of thickness 0.143 m, and relative permittivity ϵ_r = 7.6. The imaged scene is depicted in Fig. 4.7. It contains a 0.4 m high and 0.3 m wide dihedral, placed on a turntable made of two 1.2 m x 2.4 m sheets of 0.013 m thick plywood. A step-frequency signal between 0.7 and 3.1 GHz, with 3 MHz frequency step, was used to illuminate the scene.
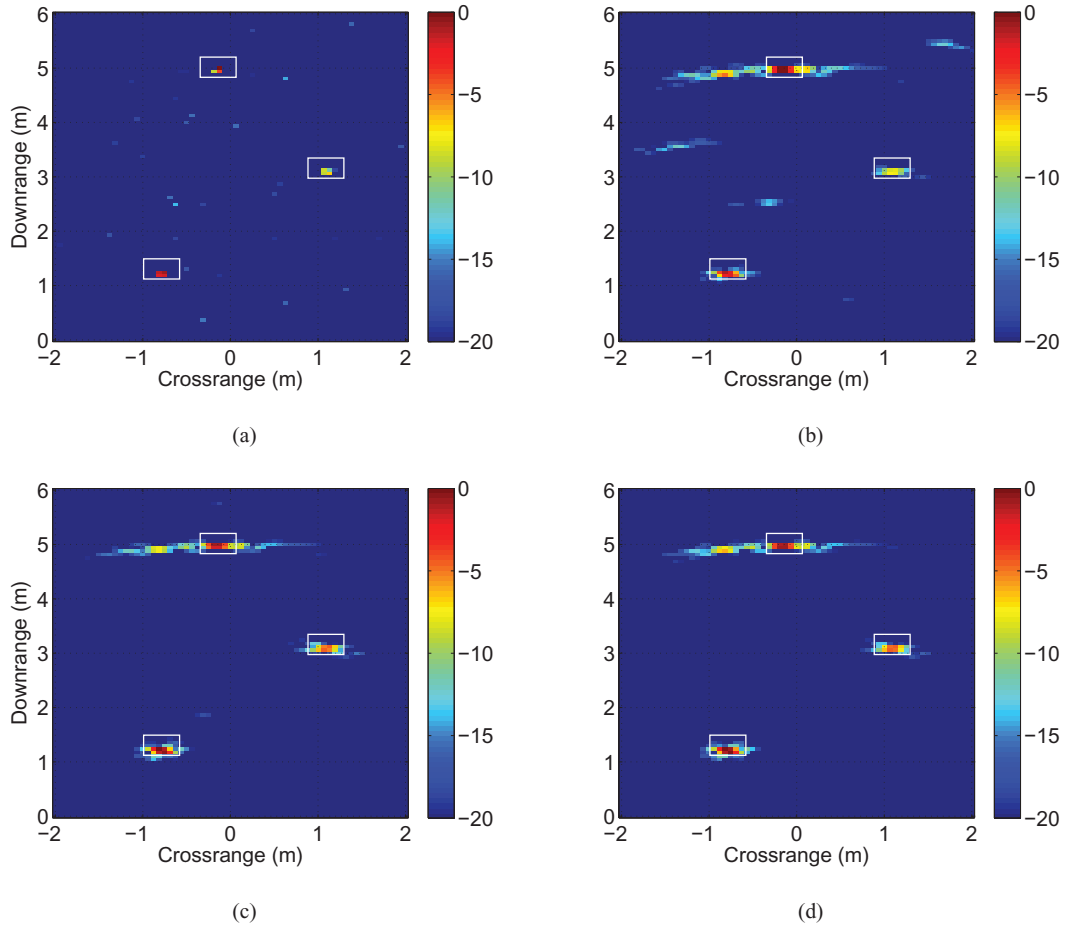
Figure 4.6: Scene images formed by different imaging approaches: (a) CS image formed by method in [89] ; (b) DS image formed by method in [91]; (c) DS image formed by the proposed approach with wavelet dictionary; (d) DS image formed by the proposed approach with Gabor dictionary. All antenna locations are used and the frequency bins are just 20% of the total transmitted frequency. The signal is corrupted by the noise with SNR = 10 dB.

The antenna array was placed at a height of 1.22 m above the floor and a standoff distance of 1.016 m away from the wall. The antenna array was 1.24 m long, with inter-element spacing of 0.022 m. Therefore, the number of antenna elements is $M = 57$ and the number of frequencies is $N = 801$; the full measurement vector $\mathbf{z}$ comprises $M \times N = 57 \times 801 = 45{,}657$ samples. The imaged scene, extending from [0, 3] m in downrange and [−1, 1] m in crossrange, is partitioned into 81×54 pixels.

To quantify the performance of the various imaging methods, we use the target-to-clutter ratio (TCR) as a measure of quality of reconstructed images. The
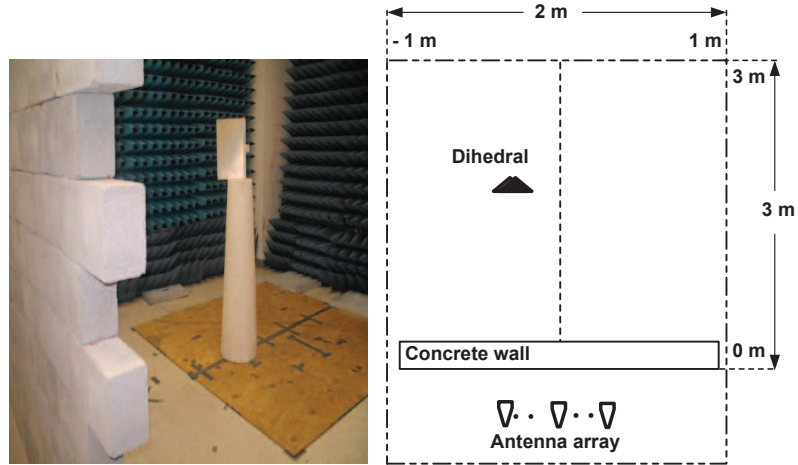
Figure 4.7: TWRI data acquisition. Left: a photo of the scene; Right: a top-view of the behind-the-wall scene.

TCR is defined as the ratio between the maximum magnitude of the target pixels and the average magnitude of clutter pixels (in dBs) [96]:

$$\text{TCR} = 20 \log_{10} \frac{\max_{(x,y) \in R_t} |I(x,y)|}{\frac{1}{N_c} \sum_{(x,y) \in R_c} |I(x,y)|}, \tag{4.19}$$

where R_t is the target area, R_c is the clutter area, and N_c is the number of pixels in the clutter region. The target region is an area of 2×6 pixels selected manually around the true target position.
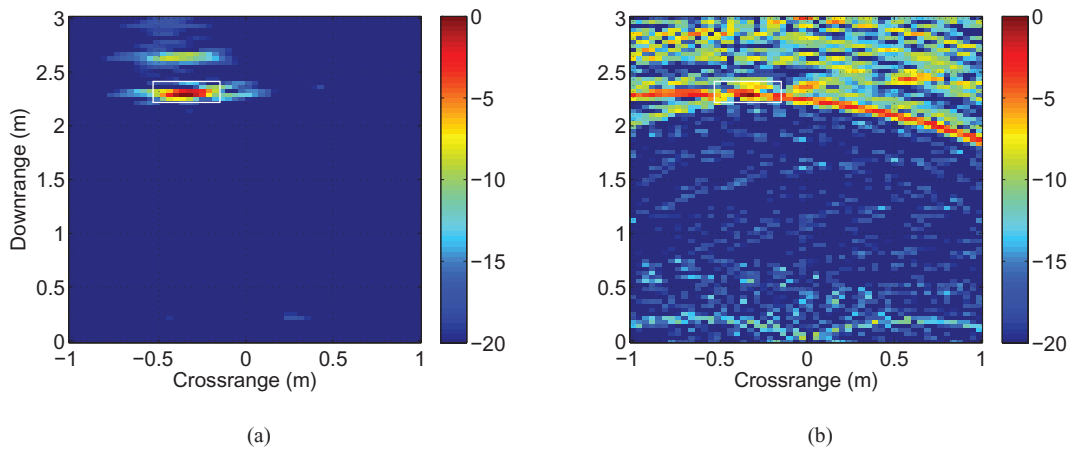
Figure 4.8: Images formed by different settings: (a) Conventional DS using full data volume; (b) conventional DS using 0.9% full data volume.

For reference purposes, Fig. 4.8(a) presents the image formed by the standard DS beamforming method using the full data volume. If all available data samples

53

are used, the conventional DS beamforming method yields a high-quality image (TCR = 30.33 dB). However, when the number of samples is significantly reduced, the standard DS beamforming method alone does not yield a high-quality image. Figure 4.8(b) shows the image formed using 2 antenna locations and 200 frequency bins (i.e., 0.9% of the collected data). Clearly this image contains too much clutter (TCR = 16.76 dB).
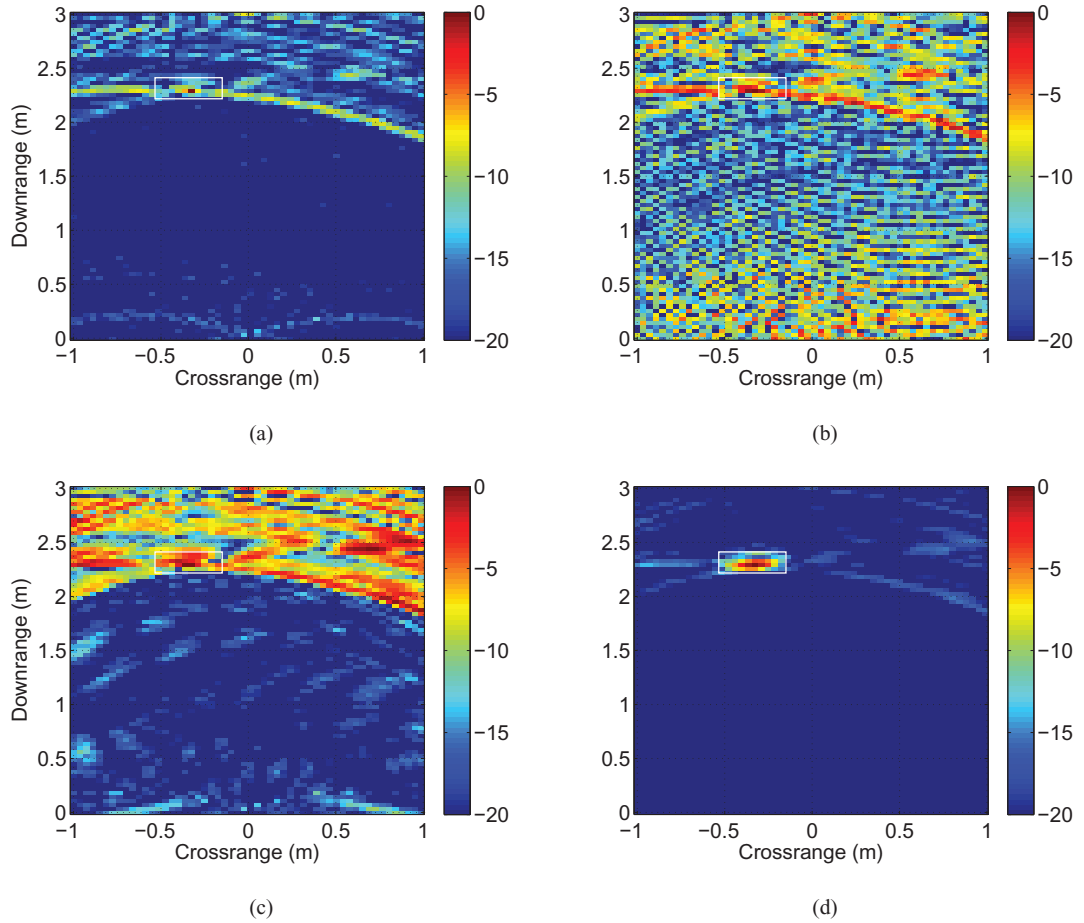
Figure 4.9: Images reconstructed by different imaging methods: (a) CS image by imaging method in [89]; (b) DS image by imaging method in [91]; (c) DS image by proposed approach without Gabor dictionary; (d) DS image by proposed approach with Gabor dictionary. The measurements constitute 0.9% of full data volume.

Using the same reduced dataset, we compare the proposed approach with other CS-based TWRI methods. Figure 4.9(a) shows the standard CS image formed using the approach in [89]. This is a significantly degraded image, compared to the image in Fig. 4.8(a), obtained using DS beamforming with full

measurements. The reason is that the imaging method in [89] directly forms the scene image by solving the conventional CS problem; when the measurements are drastically reduced and the CS solution is moderately sparse due to the presence of clutter and noise, the reconstructed image becomes less accurate. Because of the appearance of heavy clutter in Fig. 4.9(a) , the TCR of the formed image drops to 21.78 dB. Figure 4.9(b) presents the image formed by the temporal frequency

(a)   (b)
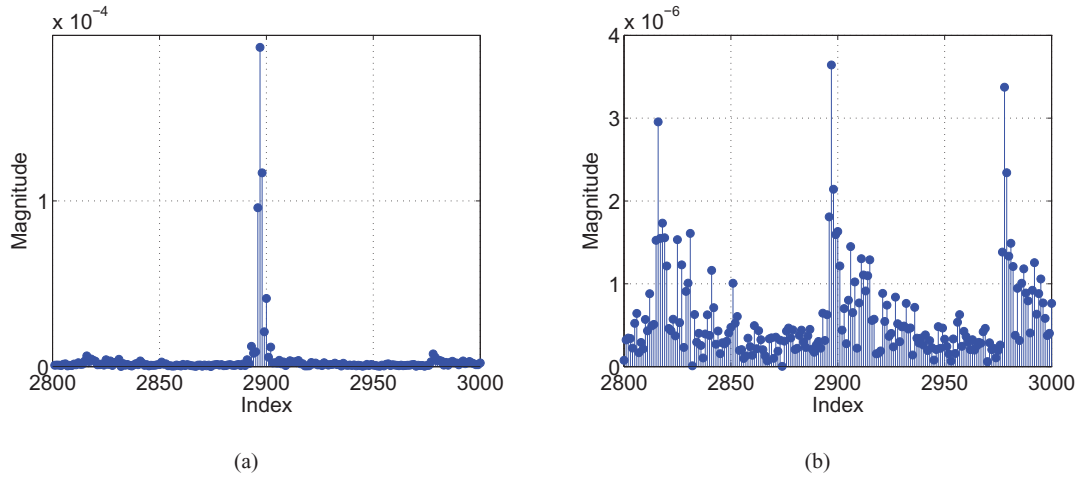
Figure 4.10: Reconstructed signal coefficients **s** for the dihedral scene: (a) using the Gabor signal representation; (b) without using the Gabor signal representation.

CS method of [91]. The quality of the formed image deteriorates because this method does not recover the full data volume when the antenna locations are reduced. The background noise and clutter appear with stronger intensity than the target in the reconstructed image (TCR = 12.13 dB). Figure 4.9(c) and 4.9(d) show, respectively, images formed by the proposed approach without and with the sparsifying Gabor dictionary. It can be observed that the image in Fig. 4.9(c), formed without the Gabor sparsifying basis, contains high clutter (TCR = 14.40 dB) and false targets. By contrast, Fig. 4.9(d) shows that the image formed using the proposed approach is considerably enhanced by incorporating the Gabor dictionary; the true target is located accurately and the clutter is considerably suppressed (TCR = 28.82 dB).

The effectiveness of the proposed approach is partly due to the excellent space-frequency localization of Gabor atoms. The Gabor functions are optimum in the
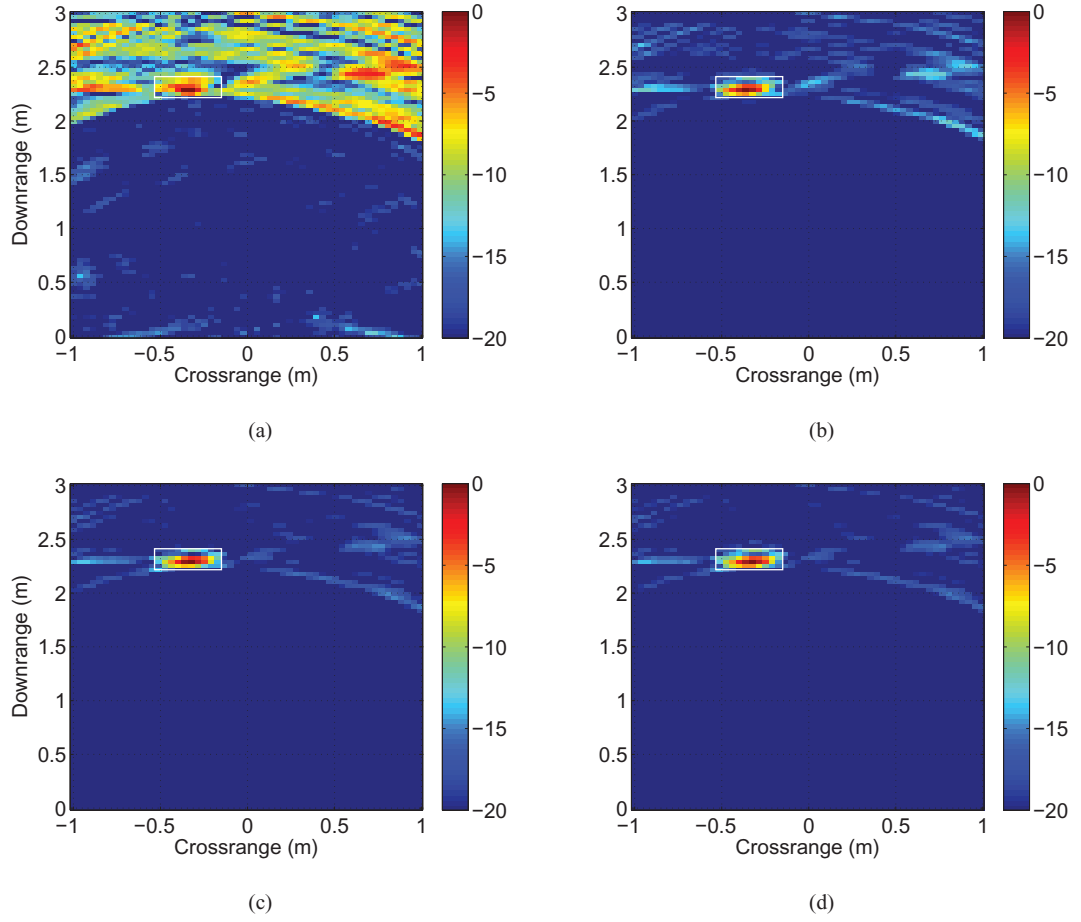
Figure 4.11: Images formed by the proposed approach with different sparsifying basis: (a) Daubechies 8 (TCR=18.56 dB); (b) Coiflet 2 (TCR=26.46 dB); (c) DT-CWT (TCR = 28.71); (d) complex Gabor dictionary (TCR = 28.82 dB).

sense that they achieve the minimum space-bandwidth product (by analogy to time-bandwidth product), which gives the best tradeoff between signal localization in space and spatial frequency domains. Figure 4.10 shows the recovered signal coefficients **s** for the dihedral scene shown in Fig. 4.7. The signal coefficients recovered with the Gabor dictionary, shown in Fig. 4.10(a), are much more sparse and concentrated on the target location, whereas the signal coefficients recovered without using the Gabor dictionary, Fig. 4.10(b), are more spread out.

In the final experiment, we use several wavelet families (Daubechies 8, Coiflet 2, and the dual-tree complex wavelet transform DT-CWT) as sparsifying basis, and compare their performances with that of the Gabor dictionary. All wavelet transforms use three decomposition levels. Figure 4.11 illustrates the images

formed using different wavelet transforms (Fig. 4.11(a)-(c)), and the image formed with the Gabor dictionary (Fig. 4.11(d)). It can be observed from the figure that the images reconstructed with the DT-CWT and the Gabor dictionaries are of superior quality than those obtained with the Daubechies and Coiflet wavelets. The formed images using the DT-CWT and the Gabor dictionary have similar TCRs of 28.71 dB and 28.82 dB, respectively. The superiority of the DT-CWT and the Gabor dictionaries can be explained by better directional selectivity and localization in space and spatial-frequency. However, we should note that the choice of the best dictionary for a specific TWRI system depends on many factors, such as the scene characteristics, target structure and the decomposition level.

4.5 Chapter summary

In this chapter, we proposed a new approach for TWRI image formation based on compressive sensing and delay-and-sum beamforming. The proposed approach requires significantly fewer number of frequency bins and antenna locations for sensing operations. This leads to a considerable reduction in data acquisition, processing time, and computational complexity, while producing TWRI images of almost the same quality as the delay-and-sum beamforming approach with full data volume. The experimental results demonstrate that the proposed approach produces images with considerably higher PSNRs, and is less sensitive to noise or the number of data samples used, compared to the standard delay-and-sum beamforming. Furthermore, experimental results on real TWRI data indicate that the proposed approach produces images with higher TCRs compared to other CS-based image formation methods. Our approach also produces images of similar TCRs compared to the DS beamforming approach that uses the full data volume. Therefore, it would be reasonable to expect that the proposed approach will enhance TWRI target detection, localization and classification, while allowing a reduction in the number of measurements and data acquisition time.

Bayesian scene reconstruction in compressed through-the-wall radar sensing

Chapter contents

Abstract

One major challenge in urban sensing applications is to detect stationary targets behind walls and inside enclosed structures. This chapter addresses the challenging problem of indoor scene reconstruction using through-the-wall radar imaging with only a small subset of frequency measurements. Although not all same frequency measurements are available at each antenna location, the proposed approach estimates the antenna signal coefficients simultaneously by exploiting signal sparsity and correlation among different antenna signals. A joint Bayesian sparse model is employed to reconstruct the antenna coefficients and to estimate the image of the scene. For scene reconstruction, a compact signal model is developed, whereby both the measurement vector and the dictionary are compressed, leading to a more efficient Bayesian sparse scene reconstruction. Furthermore, a subspace-projection technique is applied directly to the recovered antenna coefficients to suppress wall clutter and enhance image quality and target detection. The performance of the proposed method is evaluated extensively with simulated and real data. The experimental results show that the proposed approach yields significantly higher coefficient reconstruction accuracy and requires far fewer measurements for target detection and localization than does the conventional compressed sensing TWRI model, in which each single antenna signal coefficient is recovered separately.

5.1 Introduction

In urban sensing applications, through-the-wall radar imaging systems are used to detect and localize targets behind obstacles and enclosed structured [1, 2, 3, 4, 5, 6, 71, 101, 113]. Recently, it has been shown that TWRI and urban opera-

---

Part of this chapter has been published in V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Enhanced wall clutter mitigation for through-the-wall radar imaging using joint Bayesian sparse signal recovery," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Florence, Italy, May 2014, pp. 7804–7808. Partial content of this chapter has also been published in [11].

tions can be enhanced using compressed sensing theory, which enables a sparse signal to be reconstructed from under-sampled measurements [14, 15, 114]. Compressed sensing has been employed in many radar applications to relax signal sampling constraints and reduce data acquisition logistics [87, 88, 115]. In TWRI, CS techniques have been developed for scene reconstruction [7, 89, 91, 92, 95], multi-polarization [43, 116], and human motion indication [94, 117].

Reconstructing the image of the scene in the presence of wall clutter in a CS framework is a challenging but vital task for target detection and localization. In typical compressed through-the-wall radar sensing, compressive frequencies differ across antennas due to competing wireless services, intentional interferences, and deceptive radar jamming [118]. As a result, the signal phase returns vary along the sensing locations, thereby impeding direct application of wall clutter mitigation techniques [96, 106, 119, 120]. To overcome this problem, in [108] the antenna signals were reconstructed independently by employing CS at each antenna location. This CS-based recovery scheme is known as *single-signal CS model* [121] since each signal is recovered separately. In [122], instead of using the Fourier basis, the authors employed the discrete prolate spheroidal sequence (DPSS) basis to represent the signal, but the signal estimation was still performed independently along the antennas. The single-signal CS model produces a high signal reconstruction error when the frequency samples are drastically reduced at each antenna, leading to ineffective wall clutter mitigation and a significant reduction in image quality. Furthermore, at low compressed sampling rates, the recovery by the single-signal CS model can be unstable due to deficient data samples acquired at each antenna location.

In this chapter, we address the scene reconstruction problem in conjunction with wall-clutter rejection where compressed frequencies differ across the antennas. The proposed approach incorporates both sparsity and correlation structure among antenna signals into the imaging model. In TWRI, a scene is typically illuminated by several antennas in physical or synthetic aperture, placed parallel

to the wall at a standoff distance. Consequently, the signals received along the antenna locations follow the same model and exhibit significant correlations. In the transformed domain using sparsifying dictionary, the correlation among antenna signals also implies that the signal coefficients have overlapping support. We therefore consider the received radar returns as an ensemble of correlated signals, which allows us to exploit inter-signal structure and further reduce measurements. To relax signal sampling constraints, a joint sparse Bayesian framework is employed to simultaneously reconstruct all the antenna coefficients. In the Bayesian sparse model, the joint sparsity of the signals are enforced by using a shared prior and learning model hyperparameters [54, 55]. Here, the sparse Bayesian formulation is extended to deal with the complex-valued TWRI model. Once the coefficients are jointly estimated from compressed measurements, a subspace-projection technique is applied directly to the recovered coefficients to suppress the wall returns.

This chapter extends our previous work presented in [9] in several aspects: model formulation, sparse Bayesian inference, and experimental evaluation. The TWRI problem is formulated under a general CS model, where different subsets of frequencies are sensed at different antenna locations. By exploiting the inter-signal correlation, the stability and accuracy of the coefficient reconstruction can be improved significantly, thereby enhancing wall clutter mitigation and target detection and localization. Furthermore, the proposed model combines all measurement subsets to estimate all antenna signal coefficients. This allows a reduction in the sampling rate at each antenna, thereby saving data acquisition, computation, and operation cost. For a more efficient sparse scene reconstruction, principal component analysis is used to further compress the measurement vector and reduce noise; therefore, fast and more accurate Bayesian scene reconstruction can be achieved. The proposed TWRI technique is evaluated under different sensing conditions, specifically when the data samples are drastically reduced. Furthermore, a comprehensive experimental analysis and comparison with other

existing compressed TWRI models are conducted using both simulated EM data and real radar signals.

The remainder of the chapter is organized as follows. Section 5.2 briefly introduces the TWR signal model. Section 5.3 presents the compressed TWR sensing, and reviews a single-signal CS model for antenna signal recovery. Section 5.4 introduces the proposed joint Bayesian sparse scene reconstruction approach, including simultaneous antenna signal recovery, subspace-projection technique for wall clutter mitigation, and efficient Bayesian sparse scene reconstruction. Finally, Section 5.5 presents the experimental results and analysis, and Section 5.6 gives concluding remarks.

5.2 Through-the-wall radar signal model

Consider a monostatic stepped-frequency TWRI system where a transceiver is placed at several scan positions parallel to the wall to synthesize a horizontal M-element linear antenna array. The scene is interrogated by transceiving a stepped-frequency signal comprising N frequencies, equally spaced over the sensing bandwidth,

$$f_n = f_0 + n \, \Delta f, \text{ for } n = 0, 1, \ldots, N - 1, \tag{5.1}$$

where f_n is the n-th frequency and Δf is the frequency step size. Suppose that the scene contains P targets placed behind the wall. Let $z_{m,n}$ denote the n-th frequency signal received at the m-th antenna location. This signal can be expressed as

$$z_{m,n} = z_{m,n}^{wt} + v_{m,n}, \tag{5.2}$$

where $z_{m,n}^{wt}$ represents the signal due to the wall and target reflections, and $v_{m,n}$ is the noise term. The signal $z_{m,n}^{wt}$ is modeled as the superposition of the wall and

target returns,

$$z_{m,n}^{wt} = z_{m,n}^{w} + z_{m,n}^{t} = \sum_{r=1}^{R} \sigma_w \, a_r \, e^{-j2\pi f_n \tau_{m,w}^r} + \sum_{p=1}^{P} \sigma_p \, e^{-j2\pi f_n \tau_{m,p}}, \tag{5.3}$$

where $z_{m,n}^{w}$ and $z_{m,n}^{t}$ are the signals representing wall and target returns, respectively, σ_w is the reflectivity of the wall, R is the number of wall reverberations, a_r is the path loss factor associated with the r-th wall return, $\tau_{m,w}^r$ is the propagation delay of the r-th wall reverberation, σ_p is the reflectivity of the p-th target, $\tau_{m,p}$ is the round-trip travel time of the signal from the m-th antenna location to the p-th target.

Suppose the scene behind the wall is partitioned into a rectangular grid consisting of Q pixels along the crossrange and downrange. Let $s(q)$ denote a weighted indicator function defined as

$$s(q) = s^{w}(q) + s^{t}(q) \tag{5.4}$$

$$= \begin{cases} \sigma_w, & \text{if the wall occupies the } q\text{-th pixel;} \\ \sigma_p, & \text{if the } p\text{-th target resides the } q\text{-th pixel;} \\ 0, & \text{otherwise,} \end{cases} \tag{5.5}$$

where $s^{w}(q)$ and $s^{t}(q)$ represent for the wall and target components, respectively. We denote by $\mathbf{z}_m$, $\mathbf{z}_m^{wt}$, and $\boldsymbol{v}_m$ the column vectors containing, respectively, the frequency measurements $z_{m,n}$, $z_{m,n}^{wt}$, and the noise samples $v_{m,n}$ collected at the m-th antenna, see Eq. (5.2). Similarly, let $\mathbf{s}$, $\mathbf{s}^{w}$, and $\mathbf{s}^{t}$ be the lexicographically ordered column vectors containing the pixel values of the scene image. It follows from Eqs. (5.2) – (5.5) that

$$\mathbf{z}_m = \mathbf{z}_m^{wt} + \boldsymbol{v}_m, \tag{5.6}$$

$$\mathbf{z}_m^{wt} = \boldsymbol{\Psi}_m \, \mathbf{s} = \boldsymbol{\Psi}_m \, \mathbf{s}^{w} + \boldsymbol{\Psi}_m \, \mathbf{s}^{t}, \tag{5.7}$$

where $\boldsymbol{\Psi}_m$ is an $N \times Q$ matrix whose nq-th element $[\boldsymbol{\Psi}_m]_{nq} = \exp(-j2\pi f_n \tau_{m,q})$, with $\tau_{m,q}$ being the propagation delay between the m-th antenna and the q-th pixel. By concatenating the received signals at all M antennas, we can write

$$\mathbf{z} = \mathbf{z}^{wt} + \boldsymbol{\nu}, \tag{5.8}$$

$$\mathbf{z}^{wt} = \boldsymbol{\Psi}\,\mathbf{s} = \boldsymbol{\Psi}\,\mathbf{s}^w + \boldsymbol{\Psi}\,\mathbf{s}^t, \tag{5.9}$$

where $\mathbf{z} = [\mathbf{z}_0^T, \ldots, \mathbf{z}_{M-1}^T]^T$, $\mathbf{z}^{wt} = [\mathbf{z}_0^{wt\,T}, \ldots, \mathbf{z}_{M-1}^{wt}{}^T]^T$, $\boldsymbol{\Psi} = [\boldsymbol{\Psi}_0^T, \ldots, \boldsymbol{\Psi}_{M-1}{}^T]^T$, and $\boldsymbol{\nu} = [\boldsymbol{\nu}_0{}^T, \ldots, \boldsymbol{\nu}_{M-1}{}^T]^T$.

The image of the scene $\mathbf{s}$ can be recovered from (5.8)–(5.9) by applying delay-and-sum (DS) beamforming or backprojection [1, 2, 71]. Note though that the wall contributions need to be removed or suppressed before image formation; otherwise, the wall returns will dominate the target reflections, rendering target detection difficult or even impossible. If the full data volume, collected using all M antennas and all N frequencies, is available, wall mitigation techniques such as spatial filtering [96] or subspace projection [106, 119, 120] can be applied directly. However, for practical compressed TWR sensing, we have only a set of measurements collected from a subset of antennas with a reduced set of frequencies, which impedes direct application of both methods. This is because the phase errors across the antenna elements deprive the underlying assumption of spatial invariance of wall clutter. Hence, the missing frequency samples at each antenna position need to be recovered before applying wall-clutter mitigation techniques. In the following section, we briefly review the single-signal CS model that recovers the missing measurements independently at each antenna location.

5.3   Single-signal compressed sensing model

The signal model in Eqs. (5.8)–(5.9) represents the full data measurements collected at all M antenna locations and N frequencies. In compressed TWR sensing it is assumed that only a subset of measurements, collected from a reduced set of

antennas and frequencies, is available. Suppose only M_k ($M_k < M$) randomly selected antenna locations are available for data acquisition. Let $i_k \in [0, 1, \ldots, M-1]$ (for $k = 1, \ldots, M_k$) be the index of a selected antenna location. The signal $\mathbf{z}_{i_k}^{wt}$ due to the wall and target returns at the i_k-th location can be sparsely represented using a dictionary $\mathbf{W} \in \mathbb{R}^{N \times L}$ containing L ($L \geq N$) basis functions or atoms,

$$\mathbf{z}_{i_k}^{wt} = \mathbf{W}\, \boldsymbol{\theta}_{i_k}^{wt}, \tag{5.10}$$

where $\boldsymbol{\theta}_{i_k}^{wt}$ is an L-dimensional column vector of signal coefficients.

Instead of collecting all N frequency measurements, in compressed TWR sensing only N_f ($N_f \ll N$) frequency samples are acquired at the i_k-th antenna. The reduced measurement vector $\mathbf{y}_{i_k}$ of length N_f can be related to the full measurement vector $\mathbf{z}_{i_k}$ as follows:

$$\mathbf{y}_{i_k} = \boldsymbol{\Phi}_{i_k}\, \mathbf{z}_{i_k} = \mathbf{D}_{i_k}\, \boldsymbol{\theta}_{i_k}^{wt} + \tilde{\boldsymbol{v}}_{i_k}, \tag{5.11}$$

where $\boldsymbol{\Phi}_{i_k}$ is an $N_f \times N$ selection matrix containing a single non-zero element (equal to 1) in each row and each column, $\mathbf{D}_{i_k} = \boldsymbol{\Phi}_{i_k}\, \mathbf{W}$, and $\tilde{\boldsymbol{v}}_{i_k} = \boldsymbol{\Phi}_{i_k}\, \boldsymbol{v}_{i_k}$. Note that the sensing matrices $\boldsymbol{\Phi}_{i_k}$ and $\mathbf{D}_{i_k}$ differ across antenna locations because a different subset of frequencies is selected at each antenna. Now the task is to recover the signal coefficient vector $\boldsymbol{\theta}_{i_k}^{wt}$ from the reduced frequency measurement vector $\mathbf{y}_{i_k}$, see Eq. (5.11). Once $\boldsymbol{\theta}_{i_k}^{wt}$ is reconstructed, the full antenna signal $\mathbf{z}_{i_k}^{wt}$ is obtained using Eq. (5.10).

In the single-signal CS model, each coefficient vector $\boldsymbol{\theta}_{i_k}^{wt}$ is recovered independently from the corresponding measurement vector $\mathbf{y}_{i_k}$ by solving an ℓ_1 minimization problem:

$$\hat{\boldsymbol{\theta}}_{i_k}^{wt} = \arg\min_{\boldsymbol{\theta}_{i_k}^{wt}} \left\| \boldsymbol{\theta}_{i_k}^{wt} \right\|_1 \text{ subject to } \left\| \mathbf{D}_{i_k}\, \boldsymbol{\theta}_{i_k}^{wt} - \mathbf{y}_{i_k} \right\|_2 \leq \epsilon, \tag{5.12}$$

where ϵ is a noise bound. Here the signal sparsity is enforced via the ℓ_1-norm.

This single-signal CS model has been used for TWRI applications in [91, 108], and [122]. In [91] and [108], the dictionary **W** is constructed from a set of Fourier basis functions. In [122], the dictionary **W** is formed by the DPSS basis and the wall reverberations are also estimated independently at each antenna location. In both cases, however, the vector $\boldsymbol{\theta}_{i_k}^{wt}$ is recovered separately at each antenna location. These methods therefore do not consider the inter-signal correlations among the different antenna signals. In the next section, we propose a joint Bayesian sparse model for TWRI which exploits both the intra-signal sparsity and inter-signal correlations.

5.4 Joint Bayesian sparse model

In this section, we introduce a joint Bayesian sparse representation approach for efficient indoor scene reconstruction from compressed TWR measurements. In the proposed approach, a joint Bayesian CS model is first employed to simultaneously recover all antenna coefficient vectors, taking into account signal sparsity and inter-signal correlations. To segregate wall reverberations from target signals, a subspace-projection technique is applied directly to the recovered coefficients. Furthermore, a compact measurement model is developed whereby both the coefficient vector and dictionary are compressed using dimensionality reduction technique. Finally, the image of behind-the-wall scene is recovered using a Bayesian sparse approximation.

5.4.1 Joint signal coefficient estimation

In this subsection, we use a joint Bayesian sparse approximation model to simultaneously recover antenna coefficients $\boldsymbol{\theta}_{i_k}^{wt}$ from compressed measurements $\mathbf{y}_{i_k}$, see Eq. (5.11). The underlying assumption here is that all coefficient vectors are sparse, and more importantly they are correlated or jointly sparse. In other words, the significant coefficients have overlapping support. Note that several joint or simultaneous sparse approximation algorithms have been proposed which exploit correlations between signals [123, 124, 125]. However, these methods are not

suitable for the TWRI problem since they assume all compressed measurement sets are acquired using the same sensing matrix, which in TWRI context corresponds to the constraint of having the same frequency bins at different antenna locations. To relax this constraint, we propose to use the joint Bayesian sparse approximation framework.

The joint Bayesian sparse model aims to find the most likely signal coefficients $\boldsymbol{\theta}_{i_k}^{wt}$ assuming that they are jointly sparse and that the compressed measurements $\mathbf{y}_{i_k}$ are corrupted by noise with a known distribution. The joint sparsity is guaranteed by a shared prior defined on $\boldsymbol{\theta}_{i_k}^{wt}$, and the noise model is expressed using the likelihood term. Suppose that the noise vector $\tilde{\boldsymbol{\nu}}_{i_k}$ in (5.11) is distributed as $\tilde{\boldsymbol{\nu}}_{i_k} \sim \mathcal{CN}(\mathbf{0}, \beta\mathbf{I})$, with zero-mean and covariance $\beta\mathbf{I}$. Then, the likelihood of $\boldsymbol{\theta}_{i_k}^{wt}$ is a multivariate complex Gaussian function [52, 126]:

$$p(\mathbf{y}_{i_k}|\boldsymbol{\theta}_{i_k}^{wt}, \beta) = (\pi/\beta)^{-N_f} \exp\left(-\beta \left\|\mathbf{y}_{i_k} - \mathbf{D}_{i_k}\, \boldsymbol{\theta}_{i_k}^{wt}\right\|_2^2\right). \tag{5.13}$$

Note that in Eq. (5.13), the notation $\|\mathbf{x}\|_2^2$ means $\mathbf{x}^H\mathbf{x}$, with $(.)^H$ denoting the Hermitian transpose. Following the sparse Bayesian learning framework [52, 53], the joint sparsity of the coefficient vectors $\boldsymbol{\theta}_{i_k}^{wt}$ is enforced using a shared zero-mean complex Gaussian prior imposed on $\boldsymbol{\theta}_{i_k}^{wt}$,

$$\begin{aligned} p(\boldsymbol{\theta}_{i_k}^{wt}|\boldsymbol{\alpha}) &= \prod_{l=1}^{L} \mathcal{CN}(\theta_{i_k,l}|0, \alpha_l^{-1}) \\ &= \prod_{l=1}^{L} \frac{1}{\pi\alpha_l} \exp\left(-\frac{|\theta_{i_k,l}^{wt}|^2}{\alpha_l}\right) = \mathcal{CN}(\mathbf{0}, \mathbf{A}), \end{aligned} \tag{5.14}$$

where $\mathbf{A} = \mathrm{diag}(\boldsymbol{\alpha})$ and $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_L]$ is the vector of hyperparameters used to capture the correlations among coefficient vectors $\boldsymbol{\theta}_{i_k}^{wt}$. During sparse Bayesian learning, α_l diverges to infinity, then the values $\{\theta_{i_k,l}^{wt}\}_{k=1}^{M_k}$ are zero due to the zero-mean and zero-variance Gaussian prior at this location; otherwise, they are all non-zero coefficients with different scales.

Given the hyperparameters α and β, combining the likelihood (5.13) and prior (5.14), we can express the posterior of $\theta_{i_k}^{wt}$ as

$$p(\theta_{i_k}^{wt}|\mathbf{y}_{i_k}, \alpha, \beta) = \frac{p(\mathbf{y}_{i_k}|\theta_{i_k}^{wt}, \beta)\, p(\theta_{i_k}^{wt}|\alpha)}{p(\mathbf{y}_{i_k}|\alpha, \beta)} \tag{5.15}$$

$$= \frac{1}{(\pi)^{N_f}|\Sigma_{i_k}|} \exp\left\{-(\theta_{i_k} - \mu_{i_k})^H\, \Sigma_{i_k}^{-1}\, (\theta_{i_k} - \mu_{i_k})\right\}. \tag{5.16}$$

This is the multivariate complex Gaussian distribution, $\mathcal{CN}(\theta_{i_k}|\mu_{i_k}, \Sigma_{i_k})$, with the mean and covariance given by [54, 126]:

$$\mu_{i_k} = \beta\, \Sigma_{i_k}\, \mathbf{D}_{i_k}^H\, \mathbf{y}_{i_k}, \tag{5.17}$$

$$\Sigma_{i_k} = (\beta\, \mathbf{D}_{i_k}^H\, \mathbf{D}_{i_k} + \mathbf{A})^{-1}. \tag{5.18}$$

Now, the task is to estimate the hyperparameters α and β, which are obtained by maximizing the marginal likelihood, or equivalently its logarithm:

$$\mathcal{L}(\alpha, \beta) = \sum_{k=1}^{M_k} \log p(\mathbf{y}_{i_k}|\alpha, \beta), \tag{5.19}$$

$$= \sum_{k=1}^{M_k} \log \int p(\mathbf{y}_{i_k}|\theta_{i_k}, \beta)\, p(\theta_{i_k}|\alpha)\, d\theta_{i_k}, \tag{5.20}$$

$$= -\frac{1}{2}\sum_{k=1}^{M_k} \left[N_f \log 2\pi + \log |\mathbf{B}_{i_k}| + \mathbf{y}_{i_k}^H\, \mathbf{B}_{i_k}^{-1}\, \mathbf{y}_{i_k}\right], \tag{5.21}$$

with $\mathbf{B}_{i_k} = \beta^{-1}\mathbf{I} + \mathbf{D}_{i_k}\mathbf{A}^{-1}\mathbf{D}_{i_k}^H$.

An expectation-maximization algorithm is used to maximize $\mathcal{L}(\alpha, \beta)$ with respect to α and β. The algorithm is an iterative method that alternates between an expectation (E) step and a maximization (M) step. For the E-step, it computes the posterior mean in Eq. (5.17) and covariance in Eq. (5.18). For M-step, the iterative algorithm needs the update rules for the hyperparameters. These rules can be obtained by differentiating $\mathcal{L}(\alpha, \beta)$ with respect to α and β. Performing the

derivatives and setting the results to zero yield [127]

$$\hat{\alpha}_l = \frac{M_k - \alpha_l \sum_{k=1}^{M_k} \Sigma_{i_k,(l,l)}}{\sum_{k=1}^{M_k} |\mu_{i_k,l}|^2}, \tag{5.22}$$

$$\hat{\beta} = \frac{\sum_{k=1}^{M_k} \left[N_f - L + \sum_{l=1}^{L} \alpha_l \Sigma_{i_k,(l,l)} \right]}{\sum_{k=1}^{M_k} \|\mathbf{y}_{i_k} - \mathbf{D}_{i_k} \boldsymbol{\mu}_{i_k}\|_2^2}, \tag{5.23}$$

where $\mu_{i_k,l}$ is the l-th component of $\boldsymbol{\mu}_{i_k}$, and $\Sigma_{i_k,(l,l)}$ is the l-th diagonal element of Σ_{i_k}. Note that $\hat{\boldsymbol{\alpha}}$ and $\hat{\beta}$ are a function of the mean $\boldsymbol{\mu}_{i_k}$ and covariance Σ_{i_k}, while $\boldsymbol{\mu}_{i_k}$ and covariance Σ_{i_k} are a function of $\boldsymbol{\alpha}$ and β. Therefore, the expectation-maximization algorithm iterates between (5.17)-(5.18) and (5.22)-(5.23) until convergence is achieved or a stopping criterion is satisfied [52, 63]. Once the estimated hyperparameters $\hat{\boldsymbol{\alpha}}$ and $\hat{\beta}$ are obtained, the signal coefficient vector $\hat{\boldsymbol{\theta}}_{i_k}^{wt}$ is given by the mean of the posterior in Eq. (5.17):

$$\hat{\boldsymbol{\theta}}_{i_k}^{wt} = \boldsymbol{\mu}_{i_k}|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}, \beta=\hat{\beta}} = (\beta \, \Sigma_{i_k} \, \mathbf{D}_{i_k}^{H} \, \mathbf{y}_{i_k})|_{\boldsymbol{\alpha}=\hat{\boldsymbol{\alpha}}, \beta=\hat{\beta}}. \tag{5.24}$$

Given the compressed measurement sets $\{\mathbf{y}_{i_k}\}_{k=1}^{M_k}$, the iterative algorithm can be summarized by the following steps:

1. Initialize $\boldsymbol{\alpha} = \mathbf{1}$, a non-negative random initialization, and $\beta = \text{var}[\mathbf{y}_{i_k}]^{-1}$.

2. Compute the posterior mean $\boldsymbol{\mu}_{i_k}$ using (5.17) and covariance Σ_{i_k} using (5.18).

3. Update hyperparameters $\boldsymbol{\alpha}$ using (5.22) and β using (5.23).

4. Iterate Steps 2 and 3 until convergence to a fixed point $\hat{\boldsymbol{\alpha}}$.

5. Obtain the estimated signal coefficient vectors using (5.24).

Here, it is important to note that all the compressed measurement sets $\{\mathbf{y}_{i_k}\}_{k=1}^{M_k}$ are employed for the estimation of the hyperparameters, as shown by the summation of the conditional distributions in (5.19). Hence, the correlations between the antenna signals are exploited through the learning of the hyperparameters. Once

the signal coefficient vectors $\hat{\boldsymbol{\theta}}_{i_k}^{wt}$ are obtained, the scene image reconstruction can proceed. However, the estimated coefficient vectors $\hat{\boldsymbol{\theta}}_{i_k}^{wt}$ contain the wall-reverberation components, which can dominate the target signal and hinder the visibility of stationary targets in the image. Therefore, before image reconstruction, the wall reverberations need to be mitigated.

5.4.2 Wall coefficient mitigation

To mitigate wall reverberations, wall-clutter mitigation techniques can be applied to radar signals [96, 106, 119, 120], which are estimated from the recovered coefficients $\hat{\boldsymbol{\theta}}_{i_k}^{wt}$ using Eq. (5.10). Recently, it was shown that the wall reverberations can be segregated from the target components by incorporating a purging stage in which a predefined wall-clutter support (no more than 1.5m from the front wall) is used to indicate the wall coefficients [122]. However, this method may be empirical and removes also the near-range targets. Here, instead we apply a subspace projection method directly to the recovered coefficients to segregate the wall contributions from the target returns. The subspace-projection method is based on the relative strength of the wall returns compared to the behind-the-wall targets. It applies singular value decomposition to the recovered coefficient matrix to identify the wall subspace. Then orthogonal subspace projection is performed to remove the wall reflections from the recovered coefficients.

Let $\hat{\boldsymbol{\Theta}}$ be an $N \times M_k$ matrix obtained by arranging all the reconstructed antenna coefficients $\hat{\boldsymbol{\theta}}_{i_k}^{wt}$ into columns,

$$\hat{\boldsymbol{\Theta}}^{wt} = [\hat{\boldsymbol{\theta}}_1^{wt}, \ldots, \hat{\boldsymbol{\theta}}_{i_k}^{wt}, \ldots, \hat{\boldsymbol{\theta}}_{M_k}^{wt}]. \tag{5.25}$$

Using the singular value decomposition, the matrix $\hat{\boldsymbol{\Theta}}^{wt}$ can be expressed as

$$\hat{\boldsymbol{\Theta}}^{wt} = \hat{\mathbf{U}} \, \hat{\boldsymbol{\Sigma}} \, \hat{\mathbf{V}}^H, \tag{5.26}$$

where $\hat{\mathbf{U}} = [\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_N]$ and $\hat{\mathbf{V}} = [\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_{M_k}]$ are unitary matrices containing

the left and right singular vectors, respectively, and $\hat{\boldsymbol{\Sigma}}$ is a matrix containing the singular values arranged in descending order along the main diagonal.

In TWRI, the front wall returns are relatively stronger than the target reflections [120]. Let $\mathbf{P}_w$ denote the wall subspace

$$\mathbf{P}_w = \sum_{i \in \mathcal{W}} \hat{\mathbf{u}}_i \, \hat{\mathbf{v}}_i^H, \tag{5.27}$$

where $\mathcal{W}$ denotes the index set of the singular vectors spanning the wall subspace; $\mathcal{W}$ is determined using the classification technique presented in [119]. To suppress the wall reflections, the matrix $\hat{\boldsymbol{\Theta}}^{wt}$ is projected onto a subspace orthogonal to the wall subspace:

$$\tilde{\boldsymbol{\Theta}}^t = (\mathbf{I} - \mathbf{P}_w \, \mathbf{P}_w^H) \, \hat{\boldsymbol{\Theta}}^{wt}, \tag{5.28}$$

where $\mathbf{I}$ denotes the identity matrix. Now the wall-clutter free coefficients $\tilde{\boldsymbol{\Theta}}^t$ can be used to reconstruct the image of the scene.

5.4.3   Sparse Bayesian scene reconstruction

In this subsection, we first formulate a compact linear model relating the wall-clutter free coefficients to the image of the targets in the scene. Then we apply the sparse Bayesian approach to reconstruct the image of the scene.

To reduce the complexity of the problem and attenuate noise, we employ principal component analysis to formulate an efficient imaging model. Let $\tilde{\boldsymbol{\theta}}_{i_k}^t$ $(k = 1, \dots, M_k)$ denote the k-th column vector of $\tilde{\boldsymbol{\Theta}}^t$ given in (5.28). Vector $\tilde{\boldsymbol{\theta}}_{i_k}^t$ represents the target coefficients recovered at the i_k-th antenna location after wall clutter mitigation. From Eqs. (5.7) and (5.10), it follows that $\tilde{\boldsymbol{\theta}}_{i_k}^t$ is related to the image of the targets $\mathbf{s}^t$ in the scene by

$$\tilde{\boldsymbol{\theta}}_{i_k}^t = \mathbf{W}^\dagger \, \boldsymbol{\Psi}_{i_k} \, \mathbf{s}^t, \tag{5.29}$$

where † denotes the pseudo-inverse operator. In a basis approach, the coefficient

vectors $\tilde{\boldsymbol{\theta}}_{i_k}^t$ from all M_k antenna locations are stacked to form an $M_k N \times 1$ composite measurement vector, which is used to reconstruct the scene image. However, this approach produces a very high-dimensional coefficient vector and a huge dictionary of size $M_k N \times Q$, which leads to inefficient image reconstruction. Therefore, in the following we employ principal component analysis to compress the measurement vector and reduce the dictionary size, resulting in a more compact linear model for scene image formation.

Let $\mathbf{C}$ denote the covariance matrix of the measurement vectors given in (5.29). Using the eigendecomposition yields

$$\mathbf{C} = \mathbf{G}\,\boldsymbol{\Lambda}\,\mathbf{G}^H, \tag{5.30}$$

where $\mathbf{G} = [\mathbf{g}_1, \ldots, \mathbf{g}_N]$ is a matrix of the eigenvectors and $\boldsymbol{\Lambda}$ is a diagonal matrix of the eigenvalues, arranged in descending order. We define a projection matrix $\mathbf{P} = [\mathbf{g}_1, \ldots, \mathbf{g}_{N_p}]$ consisting of the first N_p ($N_p \ll N$) eigenvectors; N_p is determined by using an information theoretic criterion, such as Akaike Information Criterion (AIC) or Minimum Description Length (MDL) [128]. Applying the projection matrix $\mathbf{P}^H$ to (5.29) yields the compressed measurement vector $\tilde{\mathbf{y}}_{i_k} \in \mathbb{C}^{N_p}$,

$$\tilde{\mathbf{y}}_{i_k} = \mathbf{P}^H\,\tilde{\boldsymbol{\theta}}_{i_k}^t = \mathbf{P}^H\,\mathbf{W}^\dagger\,\boldsymbol{\Psi}_{i_k}\,\mathbf{s}^t = \tilde{\boldsymbol{\Psi}}_{i_k}\,\mathbf{s}^t, \tag{5.31}$$

where $\tilde{\boldsymbol{\Psi}}_{i_k} = \mathbf{P}^H\,\mathbf{W}^\dagger\,\boldsymbol{\Psi}_{i_k}$. Now, by concatenating the vectors $\tilde{\mathbf{y}}_{i_k}$ at all selected antenna locations into a composite vector $\tilde{\mathbf{y}} = [\tilde{\mathbf{y}}_1^T, \ldots, \tilde{\mathbf{y}}_{M_k}^T]^T \in \mathbb{C}^{M_k N_p \times 1}$ and stacking the individual compressed dictionaries into a composite dictionary $\tilde{\boldsymbol{\Psi}} = [\tilde{\boldsymbol{\Psi}}_1^T, \ldots, \tilde{\boldsymbol{\Psi}}_{M_k}^T]^T \in \mathbb{C}^{M_k N_p \times Q}$, we obtain the following compact linear model,

$$\tilde{\mathbf{y}} = \tilde{\boldsymbol{\Psi}}\,\mathbf{s}^t. \tag{5.32}$$

Given the measurement vector $\tilde{\mathbf{y}}$ and dictionary $\tilde{\boldsymbol{\Psi}}$, the scene image $\mathbf{s}^t$ can be obtained using Bayesian sparse approximation. In this Bayesian framework,

the likelihood for $\mathbf{s}^t$ is modeled as a multivariate complex Gaussian distribution, $p(\tilde{\mathbf{y}}|\mathbf{s}^t, \tilde{\beta}) \sim \mathcal{CN}(\tilde{\boldsymbol{\Psi}} \, \mathbf{s}^t, \tilde{\beta}\mathbf{I})$, with mean $\tilde{\boldsymbol{\Psi}} \, \mathbf{s}^t$ and covariance $\tilde{\beta}\mathbf{I}$. The scene sparsity is enforced using a zero-mean Gaussian prior:

$$p(\mathbf{s}^t|\tilde{\boldsymbol{\alpha}}) = \prod_{q=1}^{Q} \mathcal{CN}(s_q^t \,|\, 0, \tilde{\alpha}_q^{-1}), \tag{5.33}$$

$$= \prod_{q=1}^{Q} \frac{1}{\pi\tilde{\alpha}_q} \exp\left(-\frac{|s_q^t|^2}{\tilde{\alpha}_q}\right) = \mathcal{CN}(\mathbf{0}, \tilde{\mathbf{A}}),$$

where $\tilde{\mathbf{A}} = \mathrm{diag}(\tilde{\boldsymbol{\alpha}})$, and $\tilde{\boldsymbol{\alpha}} = [\tilde{\alpha}_1, \ldots, \tilde{\alpha}_Q]$ is the vector of hyperparameters.

Given the hyperparameters $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\beta}$, by Bayes rule, the posterior for $\mathbf{s}^t$ is a multivariate complex Gaussian distribution,

$$p(\mathbf{s}^t|\tilde{\mathbf{y}}, \tilde{\boldsymbol{\alpha}}, \tilde{\beta}) = \frac{p(\tilde{\mathbf{y}}|\mathbf{s}^t, \tilde{\beta}) \, p(\mathbf{s}^t|\tilde{\boldsymbol{\alpha}})}{p(\tilde{\mathbf{y}}|\tilde{\boldsymbol{\alpha}}, \tilde{\beta})}, \tag{5.34}$$

$$= \frac{1}{(\pi)^{M_k N_p}|\tilde{\boldsymbol{\Sigma}}|} \exp\left\{-(\mathbf{s}^t - \tilde{\boldsymbol{\mu}})^H \, \tilde{\boldsymbol{\Sigma}}^{-1} \, (\mathbf{s}^t - \tilde{\boldsymbol{\mu}})\right\}, \tag{5.35}$$

with the following mean and covariance:

$$\tilde{\boldsymbol{\mu}} = \tilde{\beta} \, \tilde{\boldsymbol{\Sigma}} \, \tilde{\boldsymbol{\Psi}}^H \, \tilde{\mathbf{y}}, \tag{5.36}$$

$$\tilde{\boldsymbol{\Sigma}} = (\tilde{\beta} \, \tilde{\boldsymbol{\Psi}}^H \, \tilde{\boldsymbol{\Psi}} + \tilde{\mathbf{A}})^{-1}. \tag{5.37}$$

The problem now becomes searching for the hyperparameters $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\beta}$, which are estimated by maximizing the marginal likelihood, or equivalently its logarithm:

$$\mathcal{L}(\tilde{\boldsymbol{\alpha}}, \tilde{\beta}) = \log p(\tilde{\mathbf{y}}|\tilde{\boldsymbol{\alpha}}, \tilde{\beta}),$$

$$= \log \int p(\tilde{\mathbf{y}}|\mathbf{s}^t, \tilde{\beta}) \, p(\mathbf{s}^t|\tilde{\boldsymbol{\alpha}}) \, d\mathbf{s}^t,$$

$$= -\frac{1}{2} \left[M_k N_p \log 2\pi + \log |\tilde{\mathbf{B}}| + \tilde{\mathbf{y}}^H \tilde{\mathbf{B}}^{-1}\tilde{\mathbf{y}}\right], \tag{5.38}$$

with $\tilde{\mathbf{B}} = \tilde{\beta}^{-1}\mathbf{I} + \tilde{\boldsymbol{\Psi}} \, \tilde{\mathbf{A}}^{-1} \, \tilde{\boldsymbol{\Psi}}^H$.

The expectation-maximization algorithm is used to maximize $\mathcal{L}(\tilde{\boldsymbol{\alpha}}, \tilde{\beta})$ with

respect to $\tilde{\alpha}$ and $\tilde{\beta}$. For the E-step, it requires the computation of the posterior mean in Eq. (5.36) and covariance in Eq. (5.37). The update rules for M-step is obtained by differentiating the cost function $\mathcal{L}(\tilde{\alpha}, \tilde{\beta})$ with respect to $\tilde{\alpha}$ and $\tilde{\beta}$. Taking the derivatives and setting the results to zero, we have the updating rules:

$$\bar{\alpha}_q = \frac{1 - \tilde{\alpha}_q \tilde{\Sigma}_{q,q}}{|\tilde{\mu}_q|^2}, \tag{5.39}$$

$$\bar{\beta} = \frac{M_k N_p - Q + \sum_{q=1}^{Q} \tilde{\alpha}_q \tilde{\Sigma}_{q,q}}{\|\tilde{\mathbf{y}} - \tilde{\boldsymbol{\Psi}} \tilde{\boldsymbol{\mu}}\|_2^2}. \tag{5.40}$$

The steps described by (5.36)-(5.37) and (5.39)-(5.40) are iterated until convergence. Then the scene image $\hat{\mathbf{s}}^t$ is given by the mean of the posterior in (5.36):

$$\hat{\mathbf{s}}^t = \tilde{\boldsymbol{\mu}}|_{\tilde{\alpha}=\bar{\alpha}, \; \tilde{\beta}=\bar{\beta}} = (\bar{\beta} \; \tilde{\boldsymbol{\Sigma}} \; \tilde{\boldsymbol{\Psi}}^H \; \tilde{\mathbf{y}})|_{\tilde{\alpha}=\bar{\alpha}, \; \tilde{\beta}=\bar{\beta}}. \tag{5.41}$$

It is important to note that the Bayesian sparse scene approximation exploits the scene sparsity, like other ℓ_1-norm minimization algorithms. However, the sparse Bayesian approach is more robust to noise and parameter settings, and it yields a more stable and precise scene, compared to conventional ℓ_1-norm solvers. This can be justified by the fact that the sparse Bayesian approach is able to learn the necessary parameters and it produces a sparse solution as a full posterior distribution, rather than a point estimate as in conventional CS solvers.

5.5 Experimental results and analysis

In this section, we present the experimental results obtained using both simulated EM data and real TWR measurements. The performance of the proposed approach is evaluated under different sensing conditions, especially when the number of measurements is drastically reduced. Experimental analysis and comparison with existing compressed TWRI models are also provided. The next subsection describes the dictionaries used for sparse representation, see Eq. (5.10), and presents briefly the performance measures used to quantify the experimental

results.

5.5.1 Dictionary selection and performance measures

Although the proposed approach works with any dictionary $\mathbf{W}$, see Eq. (5.10), in this study two sparsifying dictionaries are considered for sparse signal representation: the Fourier and the DPSS dictionary. Both dictionaries are used to compare the performance of the single-signal CS (SSCS) model [91, 108], which recovers the antenna-signal coefficients independently, with that of the proposed joint Bayesian CS (JBCS) model, which estimates the antenna-signal coefficients simultaneously. In the Fourier dictionary, the n-th column of the dictionary (for $n = 0, \ldots, N - 1$) is computed as

$$[\mathbf{W}]_n = [e^{-j2\pi f_0 \frac{2nd_r}{c}}, \ldots, e^{-j2\pi f_{N-1} \frac{2nd_r}{c}}]^T, \tag{5.42}$$

where $d_r = c/(2(N - 1)\Delta f)$ is the range resolution and c is the speed of light.

The DPSS dictionary, on the other hand, can be computed by considering each antenna signal as a multi-duration signal within the unambiguous time $[-1/(2\Delta f), 1/(2\Delta f)]$ [122]. The unambiguous time is divided into $K = \lfloor 2/(\Delta f \, \Delta T) - 1 \rfloor$ overlapping intervals of length $\Delta T = 1/(f_{N-1} - f_0)$. Let $T = (\Delta T \, \Delta f)/2$ and $t_k = [-(1/2\Delta f) + (k\Delta T/2)]\Delta f$ for $k = 1, \ldots, K$. Consider the $N \times N$ matrix $\mathbf{U} = [\mathbf{u}_0, \ldots, \mathbf{u}_{N-1}]$, consisting of the N frequency domain discrete prolate spheroidal sequences given by the eigenvectors of the matrix $\mathbf{X}$ [129],

$$[\mathbf{X}]_{i,n} = \frac{\sin(2\pi T(i - n))}{\pi(i - n)}, \text{ for } i, n = 0, \ldots, N - 1. \tag{5.43}$$

Defining an $N \times N$ diagonal matrix $\mathbf{E}_{t_k} = \text{diag}(e^{-j2\pi n t_k})$ (for $n = 0, \ldots, N-1$), we can obtain the k-th basis of the dictionary matrix $\mathbf{W}_k$ as the first $\lfloor 2NT \rfloor + 1$ columns of the matrix $\mathbf{E}_{t_k}\mathbf{U}$. By combining all the K time-shifted DPSS bases, we have the DPSS dictionary of size $N \times (\lfloor 2NT \rfloor + 1)K$: $\mathbf{W} = [\mathbf{W}_1, \ldots, \mathbf{W}_K]$.

The performance of the proposed approach is evaluated in terms of signal

reconstruction accuracy, image quality, and target detection capacity. The reconstruction error (RE) of the antenna signals is measured as

$$\text{RE} = \frac{\left\| \mathbf{z}_{i_k} - \hat{\mathbf{z}}_{i_k} \right\|_2}{\left\| \mathbf{z}_{i_k} \right\|_2}, \tag{5.44}$$

where $\mathbf{z}_{i_k}$ is the original signal and $\hat{\mathbf{z}}_{i_k}$ is the reconstructed signal at the i_k-th antenna location. For image quality, we use the target-to-clutter ratio (TCR) [113]. Let P_r denote the average power of region r in the reconstructed image I; r can be a target or clutter region. The average power P_r can be expressed as

$$P_r = \frac{1}{N_r} \sum_{(x,y) \in r} I_r^2(x, y), \tag{5.45}$$

where N_r is the number of pixels in region r. The target-to-clutter ratio is defined as the ratio between the average power of the target region and the average power of the clutter region (in dB):

$$\text{TCR} = 10 \log_{10} \left(\frac{P_{\text{target}}}{P_{\text{clutter}}} \right). \tag{5.46}$$

The ground-truth target regions are selected manually in the vicinity of the true targets, and the clutter region is the remainder of the reconstructed image.

To evaluate the performance of the imaging models in enhancing target detection, the receiver operation characteristics (ROC) curve is used to measure the probability of target detection. The probability of detection (PD), or detection rate, denotes the percentage of pixels in target regions that are correctly detected. In contrast, the probability of false alarm (PFA), or false alarm rate (FAR), is the percentage of pixels in the clutter region that are incorrectly detected as targets.

5.5.2 Simulated data

In this subsection, we evaluate the proposed approach using simulated data. Full-wave EM simulations were performed using XFDTD software package.

5.5.2.1 Simulation setup

A stepped-frequency synthetic aperture radar system was simulated for TWR data acquisition. The transceiver was placed parallel to a 0.15 m thick concrete wall, at a standoff distance of 1 m, to synthesize a linear aperture consisting of 93 elements. The spacing distance between elements was 0.02 m. A stepped-frequency signal consisting of 667 frequencies, ranging from 1 GHz to 3 GHz, with 3 MHz frequency step, was used to scan the scene. The scene layout is depicted in Fig. 5.1. It contains two dihedral targets and the front wall. The downrange and crossrange of the scene extend from 0 to 4 m, and -2 to 2 m, respectively. The pixel size is set to the Rayleigh resolution of the radar, which gives an image size of 53×36 pixels.
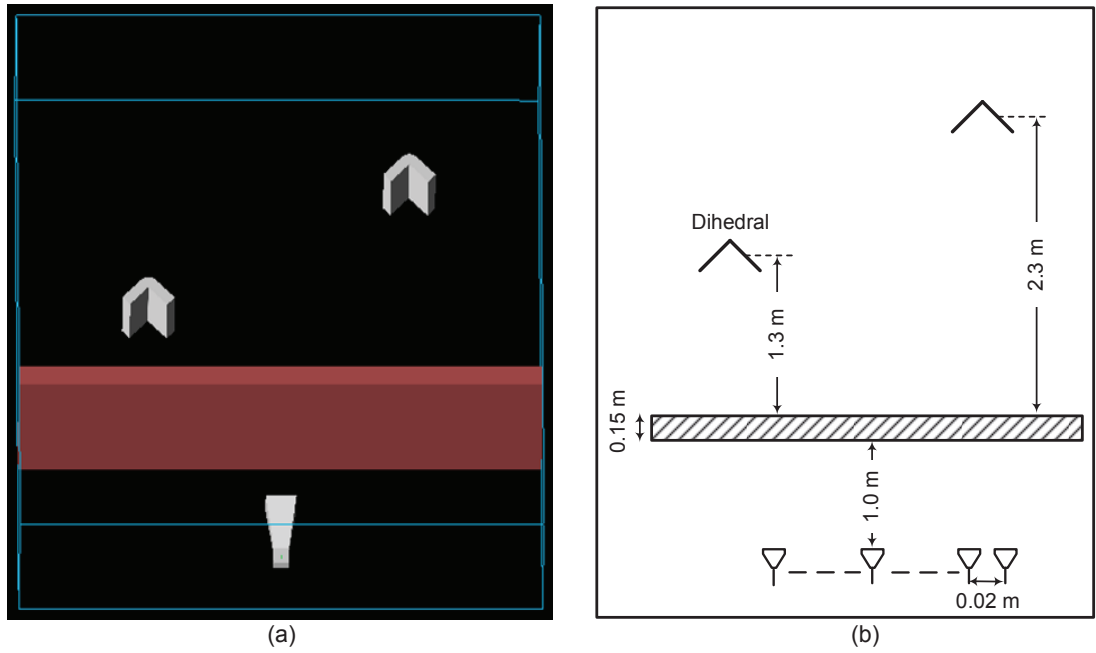
Figure 5.1: TWRI data acquisition: (a) a photo of the scene; (b) a top-view of the behind-the-wall scene.

5.5.2.2 Results and analysis

For reference purposes, Fig. 5.2 shows the scene images formed using the full data set collected using all antennas ($M = 93$) and all frequencies ($N = 667$), before wall clutter mitigation. Figure 5.2(a) presents the scene image formed by applying DS

beamforming [1] to Eqs. (5.8)–(5.9). It is clear that the wall returns dominate the target reflections, making target detection impossible. Figure 5.2(b) depicts the scene image recovered by the Bayesian sparse approximation. In this image, only the wall clutter is localized; both targets are missing. These results indicate that for target detection and localization, it is vital to mitigate wall clutter before image reconstruction. For clarity, in these and all subsequent images, wall and target regions are indicated with dashed and solid rectangles, respectively. The formed images are shown on a 50-dB scale, with the maximum intensity value normalized to 0 dB.
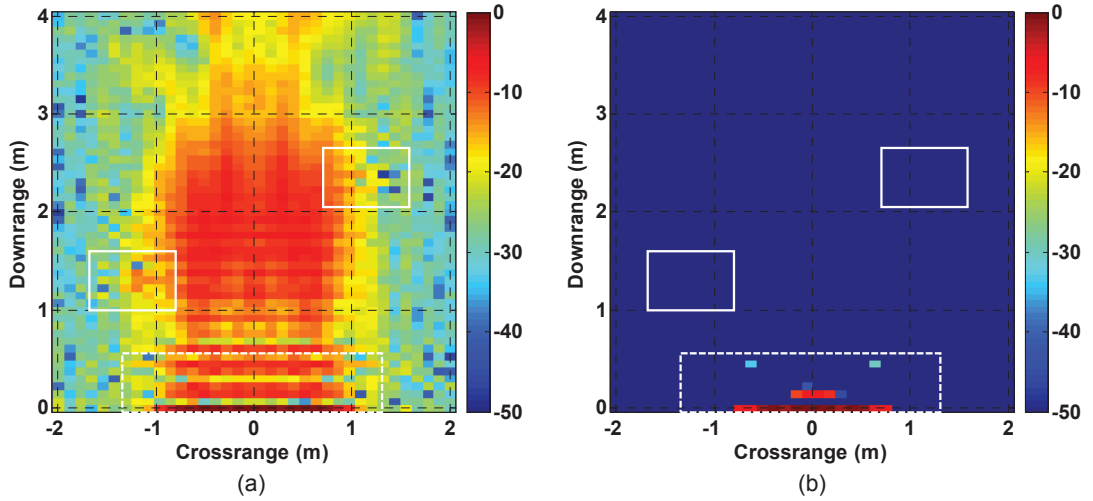
Figure 5.2: Scene images reconstructed with the full data set before wall clutter mitigation: (a) DS beamforming; (b) Bayesian scene reconstruction.

For compressed TWRI, we randomly select 20% of the antenna locations (M_k = 19 out of 93) and 10% of the frequencies (N_f = 67 out of 667), which collectively accounts for 2% of the total data volume. Here, the selected frequencies differ among the antenna locations. The antenna coefficients are recovered using the single-signal CS model and the proposed joint Bayesian CS model, followed by the sparse scene reconstruction after wall clutter mitigation, using the subspace projection method described in Subsection 5.4.2. Note that in all experiments, for the sake of fair comparison, the coefficient reconstruction problem of the SSCS model, see Eq. (5.12), is solved using the joint Bayesian approach; however, each

antenna coefficient vector is estimated independently.
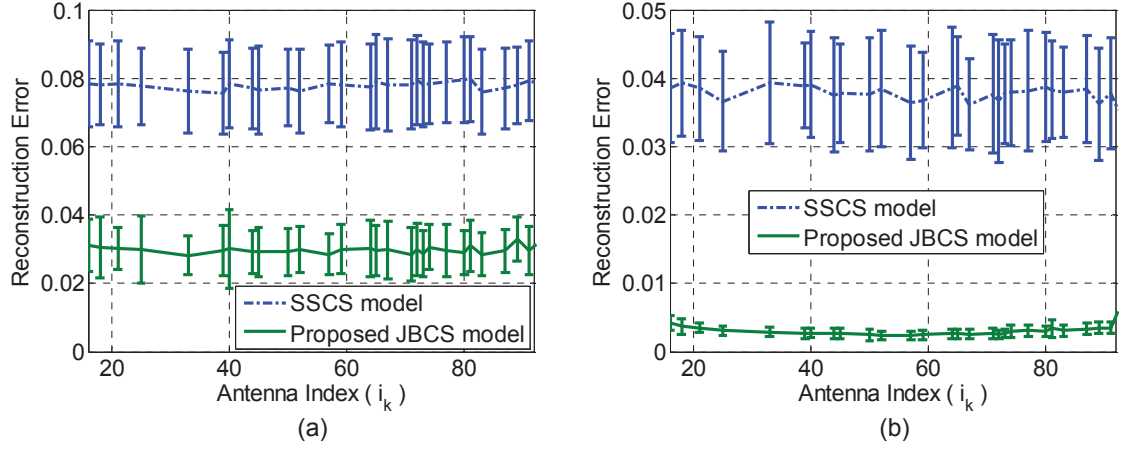
Figure 5.3: Reconstruction errors by the SSCS model [91, 108] and proposed JBCS model as a function of antenna locations: (a) using Fourier basis; (b) using DPSS basis. The results are averaged over 100 trials. The error bars represent plus/minus one standard deviation.

We conducted 100 trials and recorded the performance measures (reconstruction error, target-to-clutter ratio). Figure 5.3 shows the reconstruction errors, averaged over 100 trials, as a function of the antenna locations. Clearly, the proposed JBCS model yields significantly lower reconstruction errors than does the SSCS model. This is true for both the Fourier and DPSS dictionaries. Table 5.1 depicts the reconstruction errors by averaging along all selected antenna locations. With the Fourier basis, the proposed JBCS model yields an error of 0.029 ± 0.007, while the SSCS method gives an error of 0.078 ± 0.012. Although both models achieve a more accurate reconstruction with DPSS than with Fourier basis, the JBCS model gives a much lower reconstruction error (0.003 ± 0.001) than does the SSCS model (0.038 ± 0.008).

Table 5.1: Reconstruction performances of the SSCS model and proposed JBCS model.

Imaging models	Reconstruction Error	
	Mean	Standard Deviation
SSCS model: Fourier basis	0.078	±0.012
Proposed JBCS model: Fourier basis	**0.029**	**±0.007**
SSCS model: DPSS basis	0.038	±0.008
Proposed JBCS model: DPSS basis	**0.003**	**±0.001**
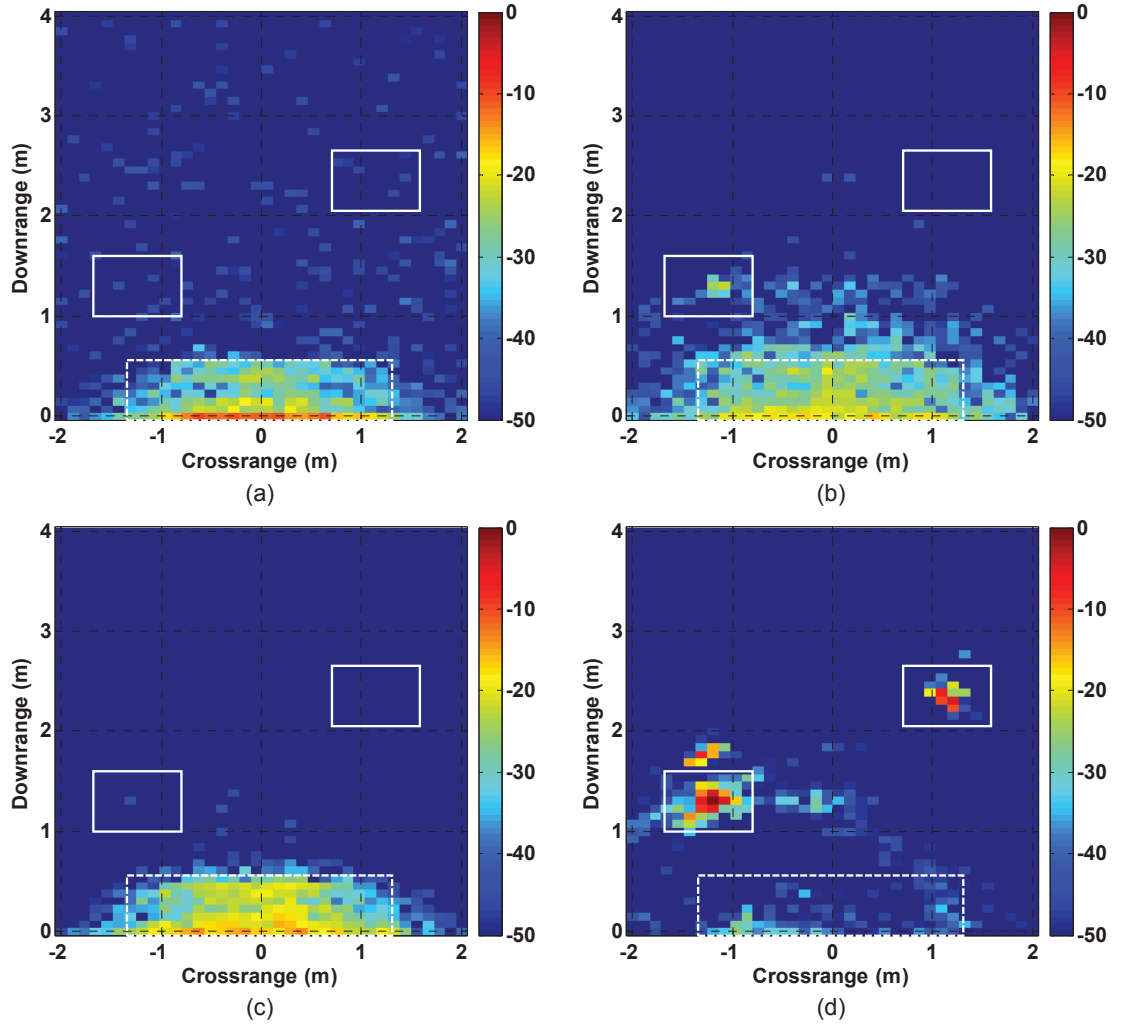
Figure 5.4: Sparse images reconstructed using 2% of the total data volume by different schemes, averaged over 100 trials; (a) SSCS model with Fourier basis; (b) proposed JBCS model with Fourier basis; (c) SSCS model with DPSS basis; (d) proposed JBCS model with DPSS basis.

Table 5.2: Average target-to-clutter ratio TCR of the scene images reconstructed by different imaging models using 2% of the full data volume.

Imaging model	TCR (dB)
SSCS model: Fourier basis [91, 108]	0.01
Proposed JBCS model: Fourier basis	**7.11**
SSCS model: DPSS basis	0.04
Proposed JBCS model: DPSS basis	**56.37**

Figures 5.4(a) and 5.4(b) show, respectively, the scene images formed using the SSCS model and the proposed JBCS model with the Fourier basis. In Fig. 5.4(a) the SSCS model fails to detect the targets, whereas in Fig. 5.4(b) the proposed JBCS model detects the dihedral target closer to the antenna array. Figures 5.4(c)

and 5.4(d) present, respectively, the images formed using the SSCS model and the proposed JBCS model with the DPSS basis. In Fig. 5.4(c), the SSCS model does not detect the targets. By contrast, the targets are clearly visible in the image of Fig. 5.4(d), formed by our JBCS model. Table 5.2 depicts the TCR values of the four images in Fig. 5.4. For both Fourier and DPSS dictionaries, there is a loss of quality in the images formed by the SSCS model, compared to the JBCS model. For example, with the DPSS basis the SSCS model achieves a TCR of 0.04 dB, whereas the proposed JBCS model yields an image with a TCR of 56.37 dB (a significant improvement of image quality).
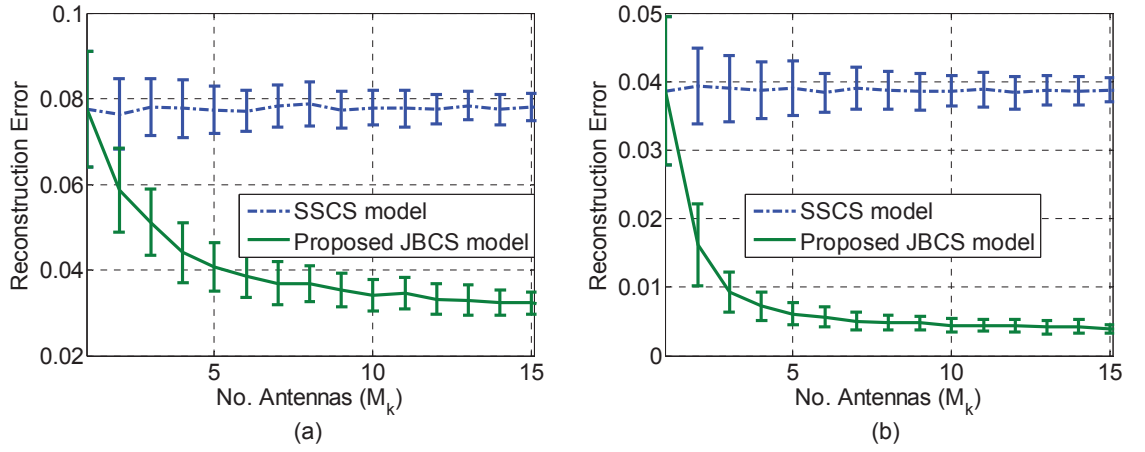
Figure 5.5: The reconstruction errors by the SSCS model and proposed JBCS model as a function of number of antennas (signals) used (M_k): (a) with Fourier basis; (b) with DPSS basis. Here 10% frequencies (N_f = 67 out of 667) are used at each selected antenna location. The results are averaged over 100 trials. The error bars represent plus/minus one standard deviation.

In the next experiment, to investigate the effects of the inter-signal correlation on the coefficient reconstruction accuracy, we fix the number of selected frequencies at 10% (N_f = 67) and vary the number of selected antennas from $M_k = 1, \ldots, 15$ (1% to 16%). For each subset of measurements, we use the SSCS model and the proposed JBCS model to reconstruct the signal-coefficients of the selected antennas, and recorded the reconstruction errors for 100 trials.

Figure 5.5 depicts the reconstruction errors by the two models as a function of the number of selected antennas M_k. We observe that for $M_k = 1$, estimating only

one coefficient vector, both models yield the same average error: 0.078±0.014 with Fourier basis and 0.038±0.01 with DPSS basis. This is because in this case the inter-signal correlation is unavailable, and thus both models exploit only the sparsity structure. However, for the other cases of reconstructing an ensemble of coefficient vectors, the average error of the SSCS model remains almost at the same value, whereas the average error of the JBCS model drops off rapidly as the number of selected antennas increases. For example, with the DPSS basis, the reconstruction error of the SSCS model remains around 0.038 ± 0.01, whereas the error value of the proposed JBCS model decreases to 0.004 ± 0.0006 (a 9.5-fold improvement). The improvement in antenna coefficient estimation can be justified by the fact that the proposed JBCS model exploits both the sparsity and inter-signal correlation, whereas the SSCS model considers only signal sparsity. As illustrated in Fig. 5.6, the antenna signals are not sparse in frequency domain, see Fig. 5.6(a). However, the signal coefficient vectors are sparse, and more importantly they share a same sparsity support, or jointly sparse in the transformed domain using a sparsifying basis, see Fig. 5.6(b).
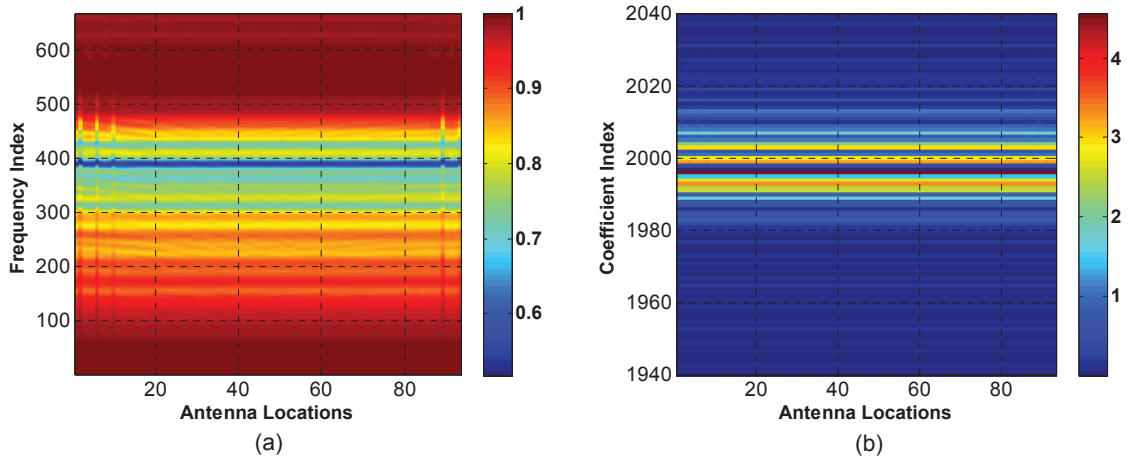
Figure 5.6: Correlations among antenna signals: (a) received signals in the frequency domain; (b) corresponding signal coefficient vectors represented using DPSS basis. Note that this figure is zoomed in on 100 significant signal coefficients only.

In the next experiment, we evaluate the performance of the proposed model under different compressed sensing ratios. We fix the number of selected antenna

locations to 20% ($M_k = 19$) and vary the ratio of selected frequencies from 5% to 20% ($N_f = 33, \ldots, 133$). Using the selected data subsets, the antenna coefficients and scene images are reconstructed by using the SSCS and JBCS models. The average performance measures, namely reconstruction error, TCR, and target detection rate, are recorded for 100 trials.
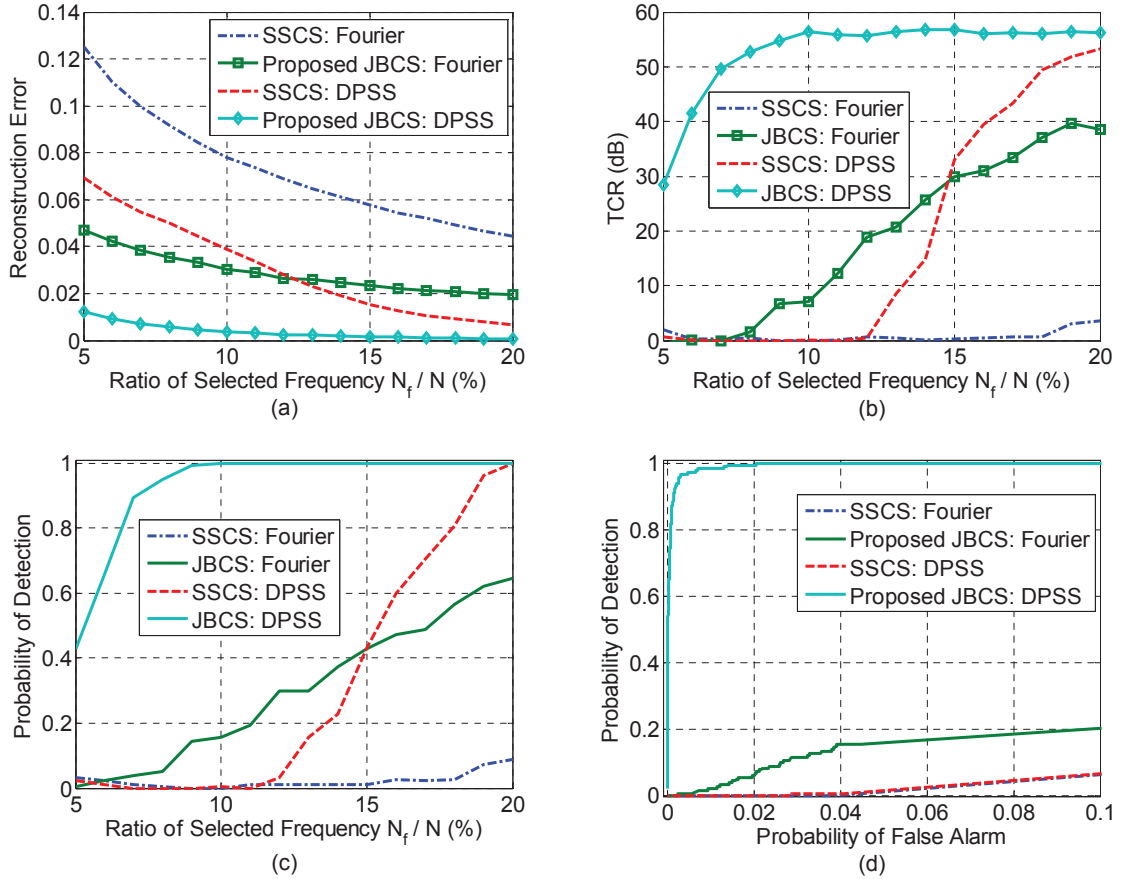
Figure 5.7: Performance comparison of the SSCS model and the proposed JBCS model: (a) the reconstruction errors as a function of percentage of selected frequencies; (b) the average TCR as a function of percentage of selected frequencies; (c) the detection rate at 5% FAR as a function of percentage of selected frequencies; (d) ROC curves at selected frequency $N_f/N = 10\%$.

Figure 5.7(a) depicts the reconstruction errors as a function of the ratio of selected frequencies to the total number of frequency measurements (N_f/N). It is clear that the proposed model produces a considerably lower error than does the SSCS model, using both dictionaries. Figure 5.7(b) shows the average TCR of the reconstructed images using both the SSCS and JBCS models. Again, the proposed JBCS model reconstructs images with substantially higher TCR than does the SSCS

model. We also found that to achieve a high-quality scene image, for example at TCR=30 dB, the SSCS model requires 15% of frequency measurements, whereas the proposed JBCS model uses only 5% with the DPSS basis.

Figure 5.7(c) presents the detection rate at 5% false alarm rate as a function of ratio of selected frequencies. From this figure we can see that the detection rate of the proposed JBCS model is much higher than that of the SSCS model; even with only 10% of selected frequency measurements, the proposed JBCS model fully detects the target (PD = 100%). In contrast, the SSCS model completely fails to detect the targets when the number of selected frequency measurements is reduced below 11%.

Figure 5.7(d) illustrates the ROC curves for 10% of selected frequencies. For a fixed false alarm rate, the JBCS improves significantly the target detection rate, in particular with the DPSS basis. Table 5.3 shows the probability of detection at 5% FAR. It is observed that the proposed approach enhances target detection significantly. For example, using DPSS basis, the proposed JBCS model has a PD of 100%, whereas the SSCS model gives a PD of 0.56%.

Table 5.3: Probability of detection at 5% false alarm rate and 10% selected frequencies (2% of full data volume).

Imaging model	Detection rate(%)
SSCS model: Fourier basis	0.0
Proposed JBCS model: Fourier basis	**15.55**
SSCS model: DPSS basis	0.56
Proposed JBCS model: DPSS basis	**100.0**

The above experimental results using simulated EM data indicate that the proposed JBCS model enhances significantly signal-coefficient estimation accuracy, and thereby yielding a substantial improvement in target detection and localization, compared to the SSCS model. We also found that using the DPSS basis is more suitable for TWRI applications, compared to the Fourier basis. This is consistent with the finding in [122]. The reason is that the DPSS functions capture the signal energy better than does the Fourier basis. In the next section, for brevity,

we use only the DPSS basis for sparse signal representation with real radar data.

5.5.3 Real data

In this subsection, the proposed approach is evaluated using real TWR measurements collected at the Radar Imaging Laboratory of the Center for Advanced Communications, Villanova University, USA. Before presenting the experimental results, we introduce the experimental setup used for simulations.

5.5.3.1 Experimental setup

Two real scenes are considered in our experiments: a dihedral scene and a calibrated scene. While the dihedral scene is suitable for sparse sensing application, the calibrated scene is more challenging since the scene sparsity is significantly reduced.

The dihedral scene layout is illustrated in Fig. 5.8. It contains a 0.4 m high by 0.3 m wide dihedral, placed on a turntable made of two 1.2 m x 2.4 m sheets of 0.013 m thick plywood. A stepped-frequency synthetic aperture radar system was used for data collection. The synthetic linear aperture consists of 57 uniformly spaced elements, with 0.022 m inter-element spacing. The antenna aperture was located at a height of 1.22 m above the floor and a standoff distance of 1.016 m away from a 0.143 m-thick solid concrete wall. The stepped-frequency signal comprises 801 frequencies in the range 0.7 to 3.1 GHz, with 3 MHz frequency step. The imaged scene extends from [0, 4] m in downrange and [−2, 2] m in crossrange. The pixel size is equal to the Rayleigh resolution of the radar, which gives an image size of 64×25 pixels.

TWR data sets were also acquired from the calibrated scene shown in Fig. 5.9. The scene contains nine targets: a sphere, a cylinder, three dihedrals, and four trihedrals. These objects were placed at different downrange, crossrange, and elevation bins. The scene was illuminated by a stepped-frequency synthetic aperture radar system. The aperture array has 69 elements, with an inter-element spacing of 0.022 m. The stepped-frequency signal comprises 201 frequencies,

equispaced over 1 GHz bandwidth centered at 2.5 GHz with 5 MHz frequency step. The radar system was placed at a distance of 0.0127 m from the wall, which was made of 0.127 m thick non-homogeneous plywood and gypsum board. The imaged scene has a downrange of [0, 8] m and a crossrange of [-3, 3] m. The pixel size is set to the Rayleigh resolution of the radar, resulting in an image of size 53×26 pixels.
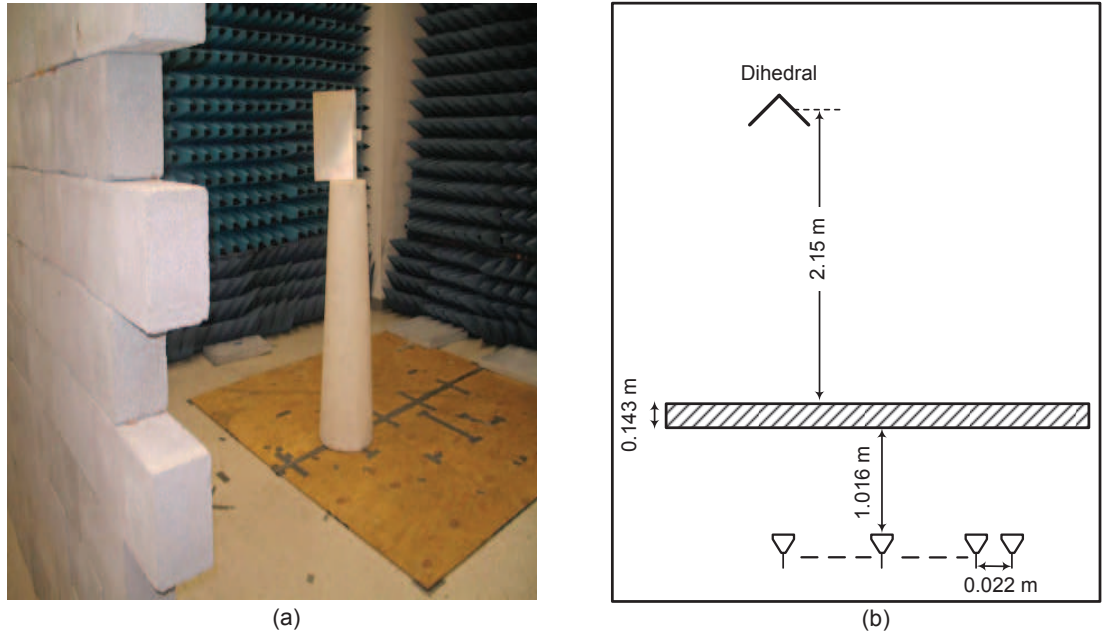
Figure 5.8: TWRI data acquisition for a dihedral scene: (a) a photo of the scene; (b) a top-view of the behind-the-wall scene.

5.5.3.2   Results for dihedral scene

For reference purposes, Fig. 5.10 shows the scene images formed using the full data set comprising all antennas ($M = 57$) and all frequencies ($N = 801$). Figure 5.10(a) presents image formed using DS beamforming, without wall clutter mitigation. It can easily be seen that the wall reverberations dominate the targets, making target detection difficult. Figure 5.10(b) depicts the scene image reconstructed using Bayesian sparse estimation. We can observe that the sparsity-based scene reconstruction recovers only the pixels in the wall region due to the wall clutter domination. Hence, for target detection and localization, it is vital to remove the wall contributions prior to scene reconstruction.
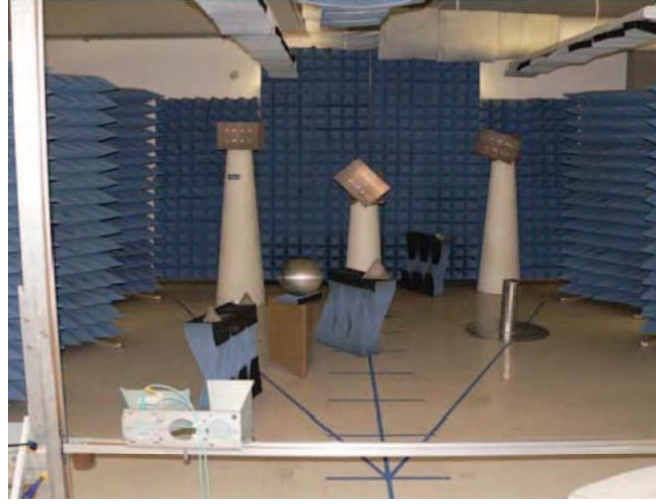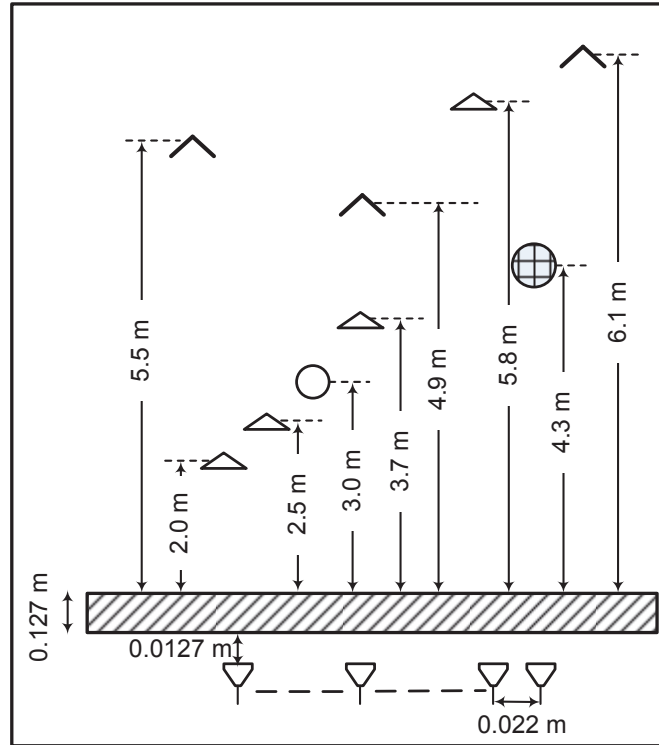
Figure 5.9: TWRI data acquisition for a calibrated scene: (a) a photo of the scene; (b) a top-view of the behind-the-wall scene.

For compressed TWR sensing, we randomly select 30% of the antenna locations ($M_k = 17$) and 10% of the frequencies ($N_f = 80$) at each selected antenna location. This collectively represents 3% of the total data volume. We conduct 100
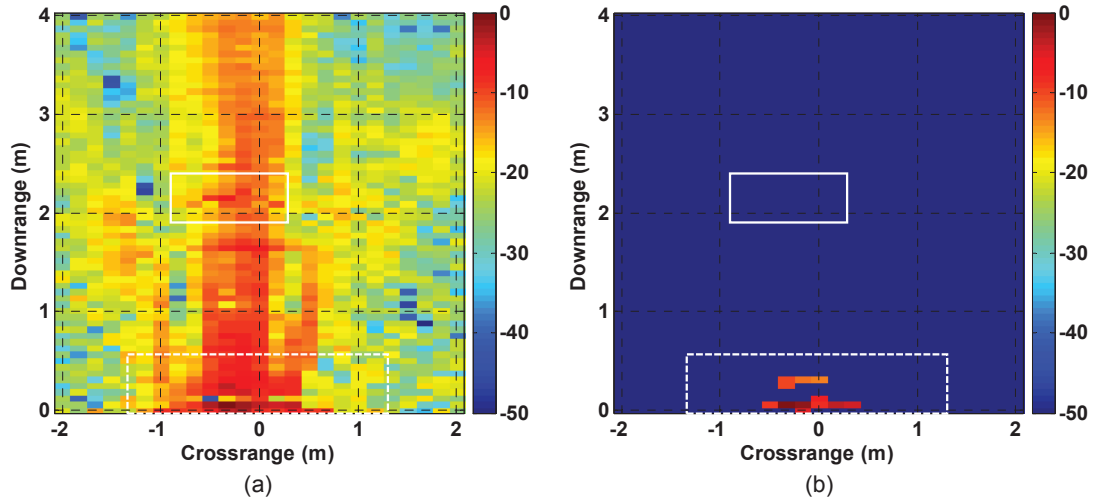
Figure 5.10: Reconstructed images using a full raw data set; (a) DS beamforming; (b) sparse scene reconstruction.

trials for antenna coefficient estimation, wall clutter mitigation, and sparse scene reconstruction. Figure 5.11 shows the formed images, averaged over 100 trials, obtained with the SSCS model and proposed JBCS model, after wall clutter mitigation using the subspace projection method. In this case, the SSCS model does not detect the target even after wall clutter mitigation, see Fig. 5.11(a), whereas the JBCS model is able to localize the target, see Fig. 5.11(b). This is expected because the proposed model yields a much higher signal-coefficient estimation accuracy, thereby achieving more effective wall clutter mitigation and image formation, compared to the SSCS model. The respective TCR values of the images in Fig. 5.11 are 1.18 dB and 42.88 dB.

Figure 5.12 shows the reconstruction errors as a function of the antenna location. The proposed model clearly produces much lower reconstruction errors than does its SSCS counterpart. Furthermore, the proposed model is much more stable than the conventional SSCS model: the standard deviation of reconstruction errors for the JBCS model is significantly smaller than that for the SSCS model. Table 5.4 presents the reconstruction errors obtained by averaging over all selected antenna locations. The proposed JBCS model improves the coefficient estimation more than 4-fold, compared with the conventional SSCS model. The stable re-
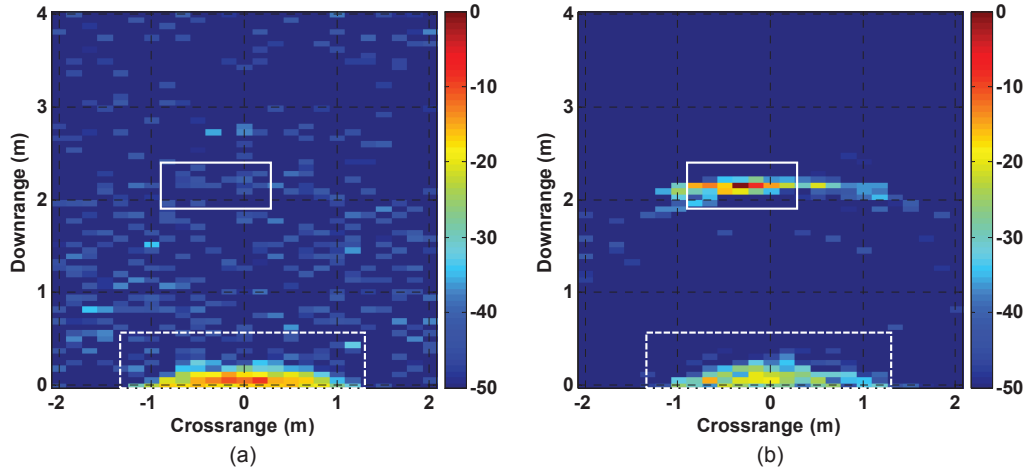
Figure 5.11: Sparse scene images reconstructed from 3% of the total data volume, averaged over 100 trials, by different schemes: (a) SSCS model; (d) proposed JBCS model.

covery is enhanced 11-fold by the proposed model, compared to the SSCS model. Fig. 5.13 shows the reconstructed coefficients at the first antenna location. Here the x-axis represents the indices of signal coefficients in the DPSS transformed domain. It is observed that the JBCS reconstruction has higher fidelity than the SSCS reconstruction.

Table 5.4: Reconstruction comparison by the SSCS model and proposed JBCS model.

Imaging models	Reconstruction Error	
	Mean	Standard Deviation
SSCS model	0.123	±0.022
Proposed JBCS model	**0.028**	**±0.002**

To evaluate the effectiveness of the models in exploiting the inter-signal correlation, we fix the selected frequency ratio at 10% ($N_f = 80$) and vary the number of selected antennas from 1% to 20% ($M_k = 1,\ldots,11$). For each reduced dataset, the antenna coefficients are reconstructed using the two models (SSCS and JBCS). The experiment is conducted for 100 trials, and the reconstruction errors are recorded. Figure 5.14 depicts the reconstruction errors as a function of the number of selected antennas M_k. This result shows that the SSCS model does not improve the reconstruction accuracy even when more measurements
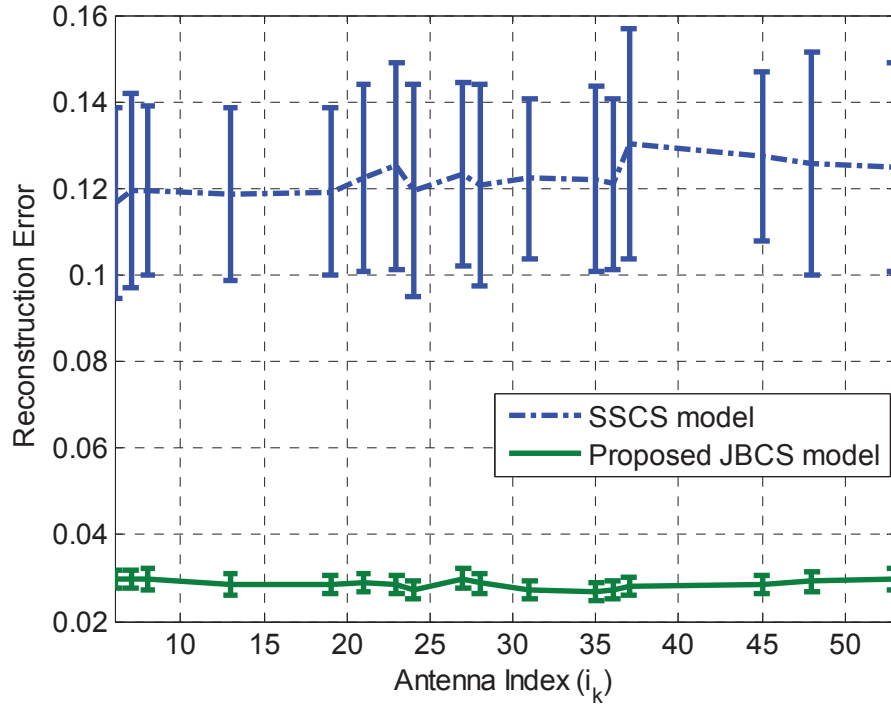
Figure 5.12: Reconstruction errors by the SSCS model and the proposed JBCS model as a function of selected antenna locations. Here, only 10% of selected frequencies are used at each antenna. The results are averaged over 100 trials. The error bars represent plus/minus one standard deviation.

are included; the reconstruction error remains almost constant at 0.123 ± 0.027. By contrast, in the proposed model, the coefficient estimation is considerably enhanced by adding more measurements from other antennas. The recovery error drops from 0.123 ± 0.027 (recovering one coefficient vector) to 0.03 ± 0.001 (recovering 11 coefficient vectors jointly). Moreover, the error bars are progressively reduced by the proposed model. In other words, the proposed model enhances both the accuracy and stability of reconstruction by exploiting the signal sparsity and inter-signal correlations.

In the next experiment, we evaluate the performance of the proposed approach as a function of the number of selected frequencies at each antenna. We set the number of selected antenna locations to $M_k = 17$ (30% of the total antennas) and vary the number of randomly selected frequencies from 40 to 240 (5% to 30% of the total number of frequencies). The experiment is conducted for 100 trials and the reconstruction errors, TCR, and PD are recorded.
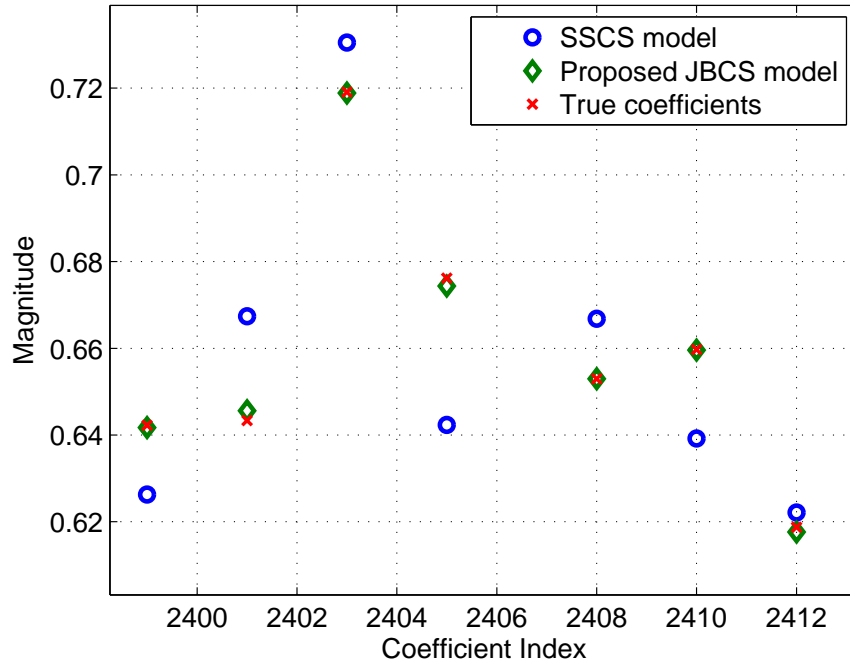
Figure 5.13: Reconstructed antenna coefficients using the SSCS model, and the proposed JBCS model. Note, this figure is zoomed in on significant coefficients only.
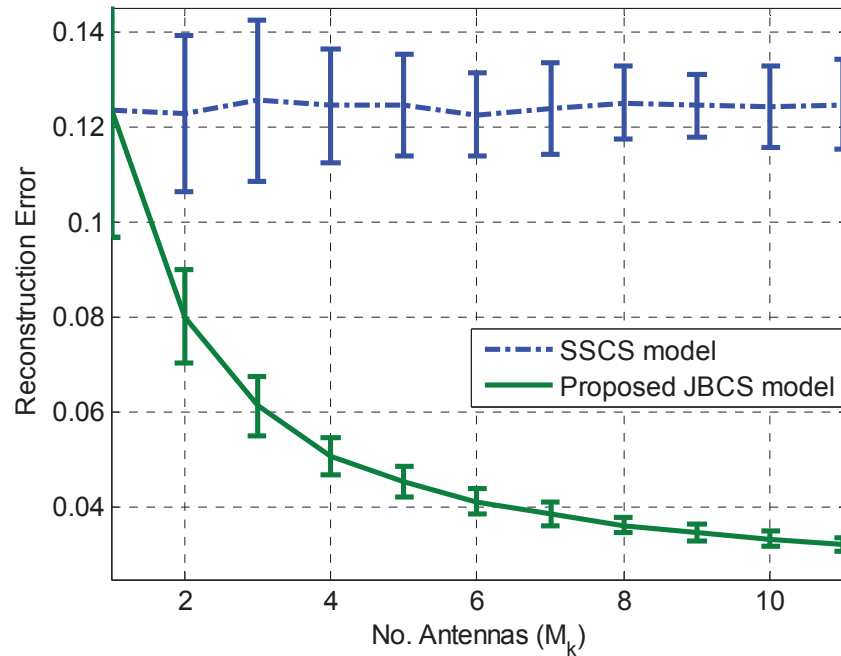
Figure 5.14: Reconstruction errors by the SSCS model and proposed JBCS model as a function of number of antenna locations used (M_k). Here, 10% frequency measurements ($N_f = 80$ out of 801) are used at each antenna. The results are averaged over 100 trials. The error bars represent plus/minus one standard deviation.
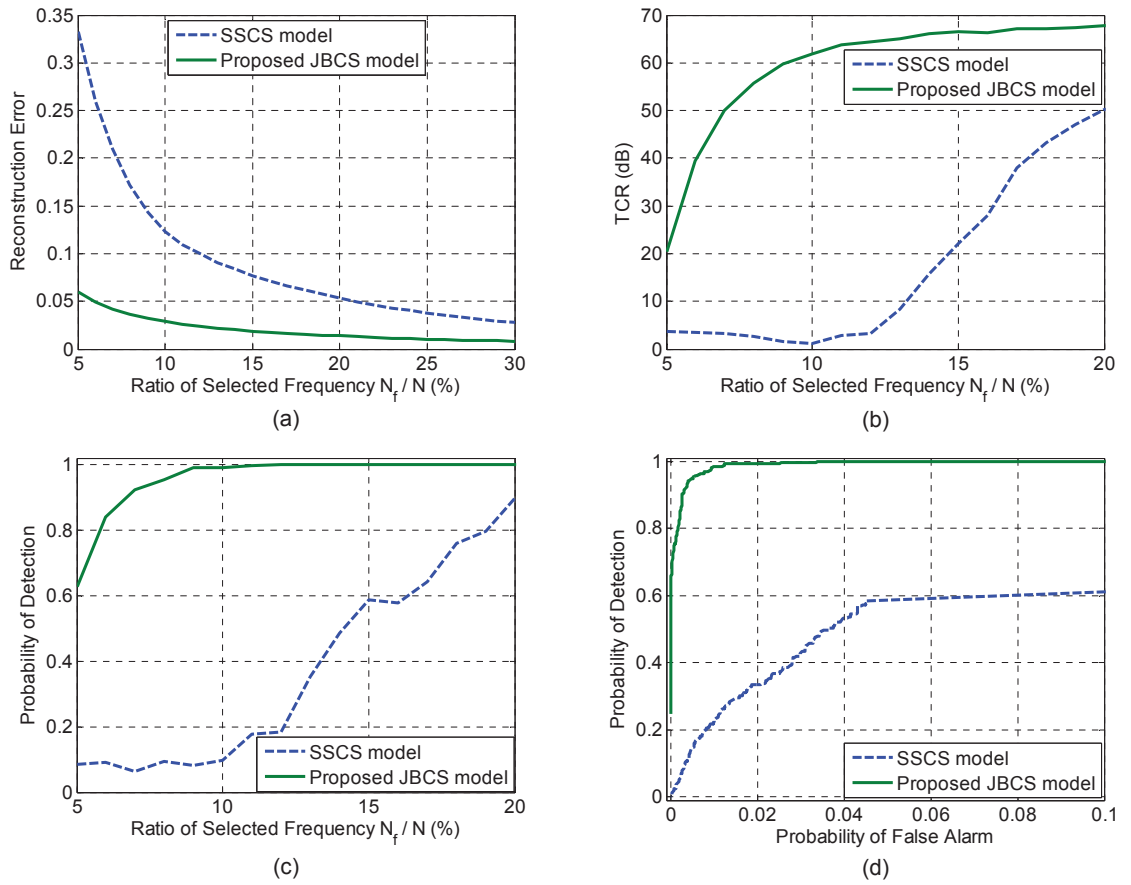
Figure 5.15: Performance comparison of the SSCS model and the proposed JBCS model: (a) the reconstruction errors as a function of percentage of selected frequencies; (b) the average TCR as a function of percentage of selected frequencies; (c) the detection rate at FAR of 5% as a function of percentage of selected frequencies; (d) ROC curves at selected frequency $N_f/N = 15\%$.

Figure 5.15 illustrates the reconstruction error, TCR, and PD as a function of the ratio of selected frequencies. Figure 5.15(a) shows that the proposed model produces a considerably lower error than does the SSCS model, especially when the measurements are reduced below 10%. Furthermore, to achieve the same reconstruction accuracy level, the proposed JBCS model requires far fewer frequency measurements than does the SSCS model. For example, to obtain an average recovery error of 0.03, the proposed model uses 10% of frequencies, whereas the SSCS model requires 30% of frequency measurements. The proposed model outperforms the SSCS model in terms of TCR, see Fig. 5.15(b). Also, the proposed model requires far fewer measurements than the SSCS model to achieve the same

TCR. For a TCR of 20 dB, the proposed model requires only 5% of the frequencies, whereas the conventional SSCS model uses 15% of the frequencies. In terms of

Table 5.5: TCR and PD at 5% false alarm rate of the formed images using 15% frequencies at each selected antennas (collectively using 4.5% of full measurements).

Imaging model	TCR(dB)	PD(%)
SSCS model	22.15	58.82
Proposed JBCS model	**66.50**	**100**
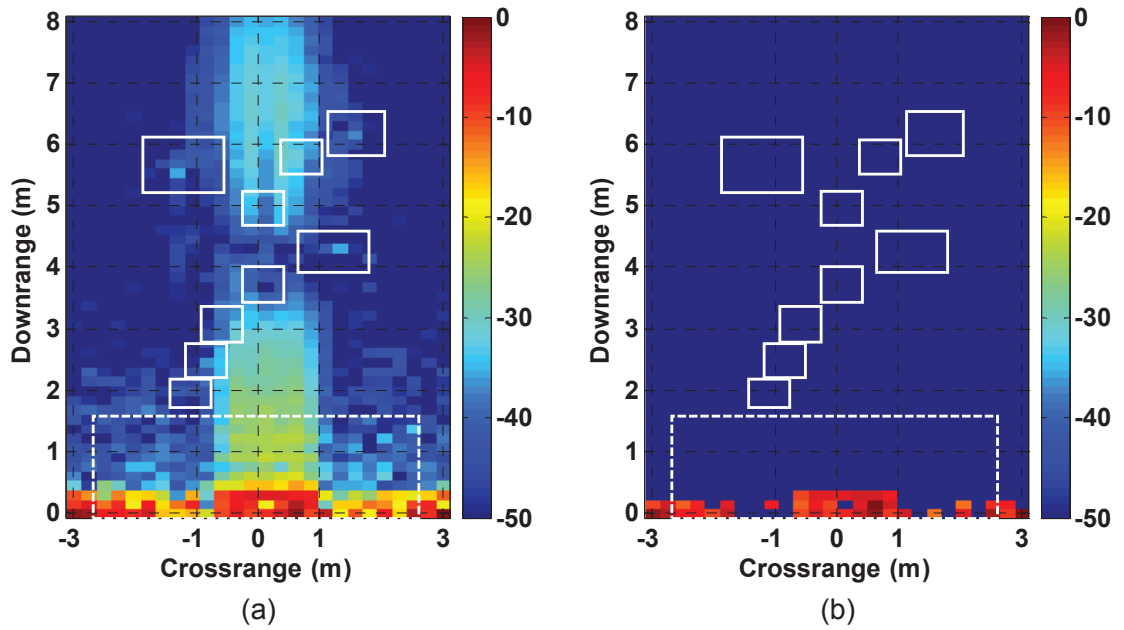
Figure 5.16: Reconstructed images using a full raw data set collected by all antennas and frequencies; (a) DS beamforming; (b) sparse scene reconstruction.

target detection, we can observe in Fig. 5.15(c) that the proposed model requires far fewer measurements for detecting target correctly, compared to its SSCS counterpart. For example, at FAR of 5%, to obtain a detection rate of 100%, the JBCS model uses only 10% of frequencies, whereas the SSCS model needs more than 20% of frequencies. The ROC curves at 15% of selected frequencies are illustrated in Fig. 5.15(d). The proposed JBCS model enhances target detection substantially. Table 5.5 depicts the TCR and PD of the two models. It is evident that the proposed JBCS model enhances image quality and target detection significantly, compared to the SSCS model.
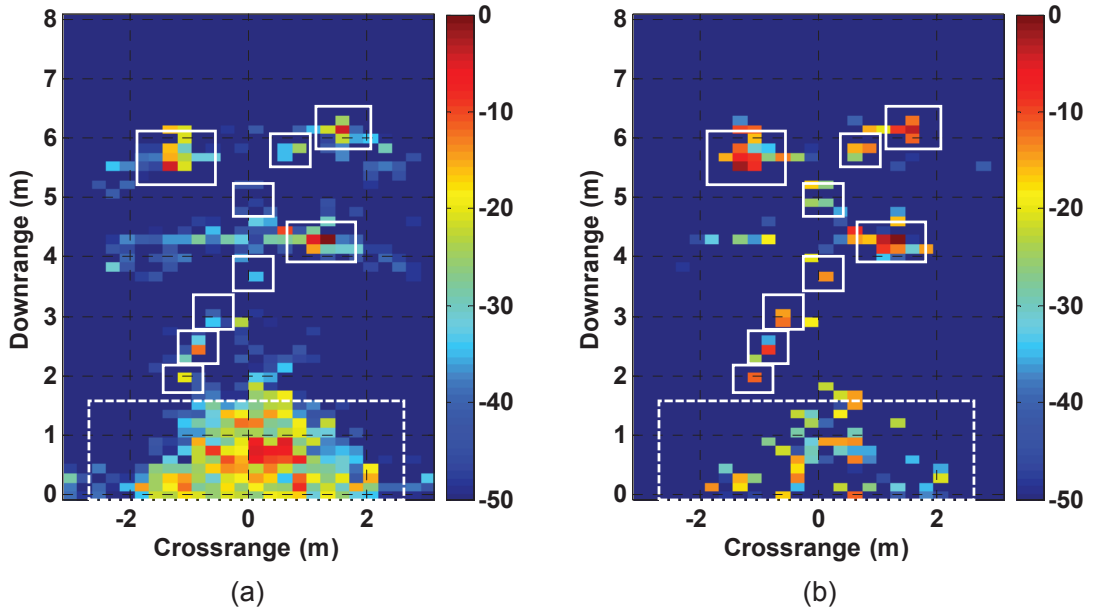
Figure 5.17: Scene images reconstructed using 30% of the full data volume: (a) SSCS model (TCR=29.06 dB); (b) proposed JBCS model (TCR=43.16 dB).

5.5.3.3 Results for calibrated scene

For reference purposes, Fig. 5.16 depicts the images formed without clutter mitigation, using the full data with all antennas ($M = 69$) and all frequencies ($N = 201$). Figure 5.16(a) presents image formed by DS beamforming. Again, the wall returns dominate the targets, impeding target detection and localization. Figure 5.16(b) shows the scene image formed by sparse Bayesian reconstruction. The sparsity-based scene reconstruction recovers only the pixels belonging to the wall reverberations. It is, therefore, necessary to remove the wall contributions prior to image reconstruction in order to detect targets behind the wall.

For compressed TWR sensing, we use full antennas ($M_k = 69$) and 30% of total frequencies ($N_f = 40$) at each selected antenna positions, which collectively represents 30% of the total data measurements. Figure 5.17 shows images, averaged over 100 trials, formed after applying the subspace projection technique to the coefficients recovered by the SSCS model and the proposed JBCS model. In the image of Fig. 5.17(a), produced by the SSCS model, the target region is much weaker than the wall region and some targets are completely missing. By
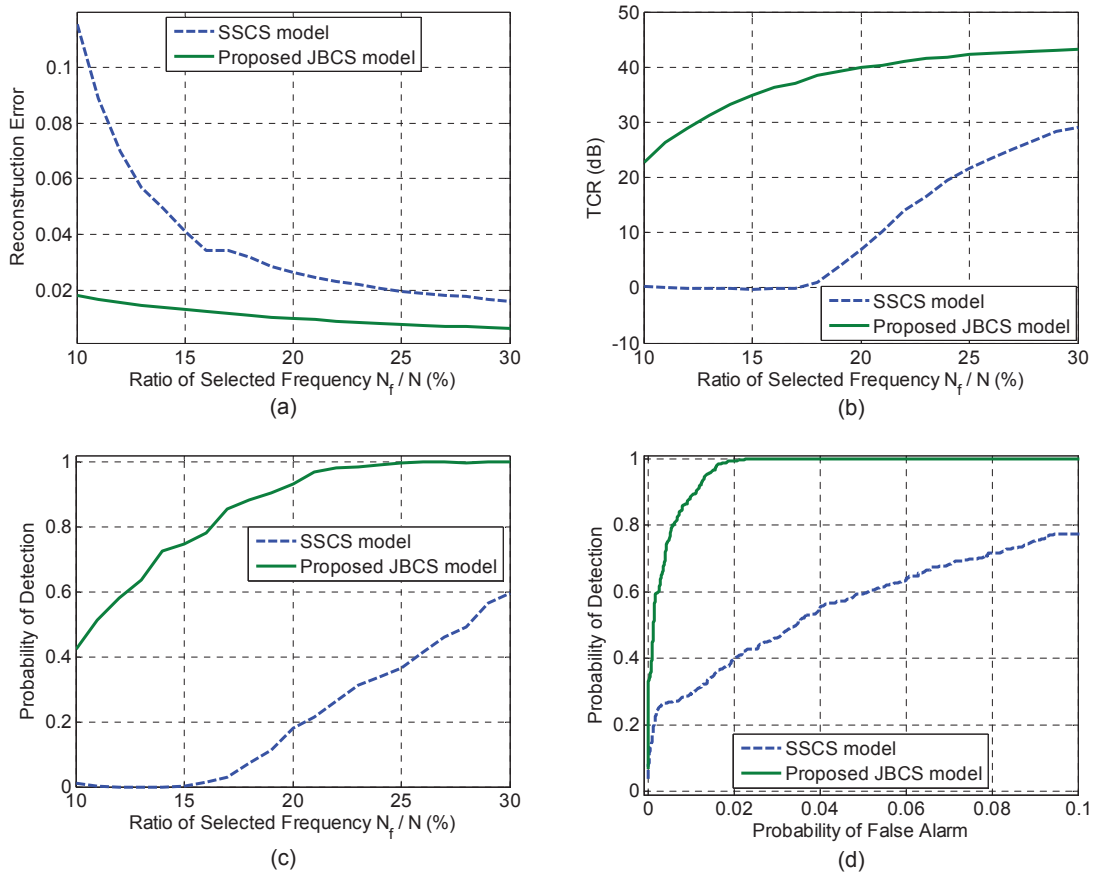
Figure 5.18: Performance comparison of the SSCS model and the proposed JBCS model: (a) the reconstruction errors as a function of percentage of selected frequencies; (b) the average TCR as a function of percentage of selected frequencies; (c) the detection rate at FAR of 5% as a function of percentage of selected frequencies; (d) ROC curves at selected frequency $N_f/N = 30\%$.

contrast, the target region in Fig. 5.17(b) is enhanced and the wall reverberations are reduced. The TCR values for these two images are, respectively, 29.06 dB and 43.16 dB.

In the last experiment, we evaluate the performance of the proposed model by varying the number of selected frequency measurements. We select frequencies from 20 to 60 (10% to 30% of the full frequencies). The antenna signal coefficients are recovered using the two models, followed with wall clutter mitigation and sparse scene reconstruction. The performance measures (reconstruction error, TCR, and PD) are recorded for 100 trials.

Figure 5.18(a) shows the reconstruction errors as a function of the ratio of the selected frequencies. It is evident that the proposed model gives a substantially

lower recovery error compared with the conventional SSCS model. For example, with 10% of selected frequencies, the reconstruction errors of the proposed approach and the SSCS model are, respectively, 0.018 and 0.115 (a 6-fold increase in signal accuracy). Figure 5.18(b) presents the average TCR of the formed images as a function of the ratio of selected frequencies. The TCR of the proposed model is consistently higher than that of the conventional SSCS model. Figure 5.18(c) depicts the PD at 5% FAR. Clearly, the proposed JBCS model outperforms the SSCS model in target detection. Figure 5.18(d) shows the ROC curves at the selected frequencies $N_f/N = 30\%$. Again, we find that the proposed JBCS model enhances target detection compared to the SSCS model.

5.6 Chapter summary

This chapter presented a joint Bayesian CS approach for indoor scene reconstruction using a reduced set of measurements collected from a subset of antennas and frequencies. It addresses the problem of compressed TWR sensing where different subsets of frequencies are sensed across different antennas. The proposed joint Bayesian sparse approximation exploits the signal sparsity structure and the inter-signal dependencies to estimate all antenna signal coefficients simultaneously. Following the joint Bayesian signal-coefficient estimation, a subspace projection technique is applied directly to the recovered coefficients to suppress the front wall radar returns. Furthermore, a compact linear imaging model is developed for efficient Bayesian sparse scene reconstruction. Extensive experimental results on simulated EM data and real radar data were presented, which demonstrate the effectiveness of the proposed imaging model, especially when the number of measurements is drastically reduced. The proposed approach enhances target detection and localization over existing compressed TWR sensing model and allows a considerable reduction in the number of measurements for high-quality indoor imaging.

Multi-channel TWRI using joint Bayesian compressed sensing

Chapter contents

Abstract

This chapter addresses the problem of scene reconstruction, incorporating wall-clutter mitigation, for compressed multi-channel through-the-wall radar imaging.

We consider the problem where the scene is sensed from multiple vantage points or different polarizations. At each channel (view or polarization), only a reduced set of antennas and frequencies are used to illuminate the same scene. The proposed approach combines wall clutter mitigation and scene reconstruction in a unified framework using multitask Bayesian compressed sensing. First, the radar signals are sparsely represented using a dictionary. Then, a joint Bayesian sparse recovery framework is employed to estimate the antenna signal coefficients simultaneously, by exploiting the sparsity and correlations between antenna signals. Furthermore, a subspace projection technique is applied to segregate the target coefficients from the wall contributions. Finally, a multitask linear model is developed to relate the target coefficients to the scene, and a composite scene image is reconstructed by a joint Bayesian sparse framework, taking into account the inter-channel dependencies. Experimental results show that the proposed approach improves reconstruction accuracy and produces a composite scene image in which the targets are enhanced and the background clutter is attenuated.

6.1 Introduction

Enhancing detection and localization of targets behind walls and inside enclosed structures using radio frequency sensors is essential for numerous civil and military applications [2, 6, 67, 130]. In practice, urban sensing and TWRI face several interferences, such as layover and shadow effects, which impede target detection and localization. For example, when the antenna is placed facing a strong reflective target with another weak target behind, layover effects occur, rendering the detection of the weak target more difficult, or impossible. Further, the target reflectivity depends highly on the sensing aspect angle. Target reflections may be strong if sensed from the front wall, but may be weak when illuminated from

---

Part of this chapter has been published in V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Multi-view indoor scene reconstruction from compressed through-wall radar measurements using a joint Bayesian sparse representation," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Brisbane, Australia, April 2015, pp. 2419–2423. Partial content of this chapter has also been published in [10, 12].

the side wall, and vice versa. These problems can be addressed by using multi-view or multi-polarization sensing and then combining the data acquired from different channels to enhance target detection and localization.

Multi-view TWRI methods typically involve image formation at individual views, followed by image fusion [6, 100], target image correction [71], or target detection [5]. These existing methods, however, are not concerned with the TWRI problem in the compressed sensing (CS) context [14, 15]. The full data volume at each view is required to form the images. In the past decade, CS has been used for TWRI to save data acquisition, reduce computation cost, and improve image formation and fusion [7, 91, 95, 116]. More recently, CS-based techniques have been proposed which combine wall-clutter mitigation with image formation [9, 108, 122]. These methods, however, are suitable for single-view TWRI problem only; they do not consider the inter-view correlations in the imaging model.

During the sensing process, the scene can be imaged from the same viewing angle, but with different polarizations. The multi-polarization approach is motivated by the fact that the strength of target returns depends on its radar cross section (RCS) — a polarimetric-dependent factor [131]. As a result, multi-polarimetric sensing provides more information about the target than does the single polarization [132, 133, 134]. Several TWRI studies have used multi-polarization of electromagnetic waves to improve the discrimination of targets [6, 102, 113, 116, 131, 135, 136, 137]. In [131], a method for TWR detection of certain types of weapons was developed by analyzing the target signature in different polarizations. In [135], a polarimetric beamforming algorithm based on Greens function was proposed for TWRI. The formed images show that the targets exhibit different characteristics with respect to the polarizations, and thus these features can be used for image enhancement. In [136], an adaptive target detection technique using multi-polarization images was developed. The technique exploits the polarization diversity to improve target detection. In [102], features extracted from multi-polarization images were employed for target classification.

In [137], a fully polarimetric scattering model of the human body was proposed for detecting a person behind walls. In [113] and [6], fuzzy logic based fusion techniques were developed to combine multi-polarization images. In [116], a multi-polarization image formation method was developed for TWRI.

Most multi-polarization TWRI studies assume that the wall reflections are removed, or at least significantly suppressed, prior to image formation. Furthermore, the existing methods assume that the same set of frequencies are collected across the antenna array aperture and polarizations, and thereby not addressing the logistic difficulties in data acquisition. The problem of imaging with multiple polarizations is that the target may exhibit different responses when interrogated by different polarized signals. Furthermore, in practical applications, not all radar signals from the antenna array aperture can be acquired successfully due to occlusions by metallic obstacles or EM interferences. Therefore, the main aim of a practical multi-polarimetric TWRI system is to effectively reconstruct scene images that have low background clutter and high target reflections from reduced polarimetric measurements.

In this chapter, we propose a new approach for addressing the problem of wall-clutter mitigation and compressed multi-channel TWR scene reconstruction using multitask Bayesian compressed sensing framework. The multi-channel sensing operation refers to the multi-view or multi-polarization sensing mode where several data sets are obtained by illuminating the same scene. In the proposed approach, a joint Bayesian sparse model is employed to reconstruct the antenna signal coefficients simultaneously, by exploiting both the sparsity and correlations between antenna signals. This joint model differs from the single-signal CS recovery model presented in [108, 122], where each antenna signal is recovered independently. This chapter demonstrates that joint reconstruction requires far fewer measurements and yields higher recovery accuracy than the single-signal CS model, where each antenna signal is reconstructed separately. Furthermore, a subspace projection technique is applied directly to the estimated

signal coefficients to segregate the wall reflections from target returns. For scene reconstruction, a multitask linear model is developed in which a fused vector of clutter-free signal coefficients is used to recover a composite scene image. The composite image of the scene and the images related to different channels are jointly reconstructed using sparse Bayesian learning framework, taking into account inter-channel dependences.

The remainder of the chapter is organized as follows. Section 6.2 introduces the multi-channel TWRI signal model. Section 6.3 describes the proposed approach, including joint Bayesian antenna signal coefficient estimation, wall-clutter mitigation, and joint Bayesian image reconstruction. Section 6.4 presents experimental results and analysis. Section 6.5 gives concluding remarks.

6.2 Multi-channel TWRI signal model

Consider a monostatic multi-channel TWRI system illuminating a scene behind a wall or inside an enclosed structure. Assume that the scene containing P targets is imaged by L channels. In multi-view mode, these L sensing views are performed by shifting the same antenna array to new locations vertically or horizontally along the front and side walls. In multi-polarimetric radar, the scene is imaged from a single view, but using different polarizations. At each channel, the TWRI system uses M antenna locations and N narrowband signals to scan the scene. In the stepped-frequency approach, a stepped-frequency signal is used consisting of N frequencies, equispaced over the desired bandwidth $f_{N-1} - f_0$,

$$f_n = f_0 + (n-1)\Delta f, \ n = 1, \ldots, N-1, \tag{6.1}$$

where f_0 is the lowest frequency in the frequency band and Δf is the frequency step size.

Let $z_l(m,n)$ denote the signal of frequency f_n, received by the m-th antenna

from the l-th channel. This signal can be expressed as

$$z_l(m, n) = z_l^{wt}(m, n) + v_l(m, n), \tag{6.2}$$

where $z_l^{wt}(m, n)$ represents the signal due to the wall and target reflections, and $v_l(m, n)$ is the noise term. The signal $z_l^{wt}(m, n)$ is modeled as the superposition of the wall and target returns,

$$z_l^{wt}(m, n) = z_l^w(m, n) + z_l^t(m, n), \tag{6.3}$$

$$= \sum_{r=1}^{R} \sigma_w a_r \exp(-j2\pi f_n \tau_{m,w}^r) + \sum_{p=1}^{P} \sigma_p^l \exp(-j2\pi f_n \tau_{m,p}), \tag{6.4}$$

where $z_l^w(m, n)$ and $z_l^t(m, n)$ are the signals representing wall and target returns, respectively, σ_w is the reflectivity of the wall, R is the number of wall reverberations, a_r is the path loss factor associated with the r-th wall return, $\tau_{m,w}^r$ is the propagation delay of the r-th wall reverberation, σ_p^l is the complex reflectivity of the p-th target at the l-the channel, and $\tau_{m,p}$ is the round-trip travel time of the signal from the m-th antenna to the p-th target.

Assume that the scene is partitioned into a rectangular grid consisting of Q pixels. Let $s_l(q)$ denote a weighted indicator function representing the scene image at the l-th channel. The weighted indicator function can be defined as

$$s_l(q) = s_l^w(q) + s_l^t(q) = \begin{cases} \sigma_w, & \text{if the wall occupies the } q\text{-th pixel;} \\ \sigma_p^l, & \text{if the } p\text{-th target occupies the } q\text{-th pixel;} \\ 0, & \text{otherwise,} \end{cases} \tag{6.5}$$

where $s_l^w(q)$ and $s_l^t(q)$ represent the wall and target components, respectively. We denote by $\mathbf{z}_{l,m}$, $\mathbf{z}_{l,m}^{wt}$, and $\mathbf{v}_{l,m}$ the column vectors containing, respectively, the frequency measurements $z_l(m, n)$, $z_l^{wt}(m, n)$, and $v_l(m, n)$ collected by the m-th antenna at the l-th channel, see Eq. (6.2). Similarly, let $\mathbf{s}_l$, $\mathbf{s}_l^w$, and $\mathbf{s}_l^t$ be the lexicographically

ordered column vectors containing the pixel values of the l-th channel image. It follows from Eqs. (6.2)–(6.5) that

$$\mathbf{z}_{l,m} = \mathbf{z}_{l,m}^{wt} + \boldsymbol{\nu}_{l,m}, \tag{6.6}$$

$$\mathbf{z}_{l,m}^{wt} = \boldsymbol{\Psi}_{l,m} \, \mathbf{s}_l = \boldsymbol{\Psi}_{l,m} \, \mathbf{s}_l^w + \boldsymbol{\Psi}_{l,m} \, \mathbf{s}_l^t, \tag{6.7}$$

where $\boldsymbol{\Psi}_{l,m}$ is an $N \times Q$ matrix. The (n,q)-th element, $\Psi_{l,m}(n,q)$, of this matrix is given by $\Psi_{l,m}(n,q) = e^{-j2\pi f_n \tau_{m,q}}$, with $\tau_{m,q}$ being the propagation delay between the m-th antenna and the q-th pixel. By concatenating the received signals at all M antennas, we can write

$$\mathbf{z}_l = \mathbf{z}_l^{wt} + \boldsymbol{\nu}_l, \tag{6.8}$$

$$\mathbf{z}_l^{wt} = \boldsymbol{\Psi}_l \, \mathbf{s}_l, \tag{6.9}$$

where $\mathbf{z}_l = [\mathbf{z}_{l,0}^T, \ldots, \mathbf{z}_{l,M-1}^T]^T$, $\mathbf{z}_l^{wt} = [\mathbf{z}_{l,0}^{wt\,T}, \ldots, \mathbf{z}_{l,M-1}^{wt\,T}]^T$, $\boldsymbol{\Psi}_l = [\boldsymbol{\Psi}_{l,0}^T, \ldots, \boldsymbol{\Psi}_{l,M-1}^T]^T$, and $\boldsymbol{\nu}_l = [\boldsymbol{\nu}_{l,0}^T, \ldots, \boldsymbol{\nu}_{l,M-1}^T]^T$.

The image of the scene $\mathbf{s}_l$ can be recovered from (6.8)–(6.9) by applying the delay-and-sum (DS) beamforming or backprojection [1, 2, 66]. However, this approach is suitable for single-channel TWRI only where the image at each channel is reconstructed independently, ignoring the correlations between channels. Note that before image formation, the wall contributions need to be removed or significantly reduced; otherwise, the wall returns dominate the target reflections, rendering target detection difficult or even impossible. In the next section, we present a new approach for compressed multi-channel TWRI which incorporates wall clutter mitigation and takes into account the correlations between antenna signals and inter-channel dependencies.

6.3 Joint Bayesian multi-channel TWRI model

This section presents the proposed approach for compressed multi-channel TWRI. In the proposed approach, the antenna signals are first represented by a sparsifying dictionary. Then, the signal coefficients are simultaneously estimated using a joint Bayesian sparse framework. Next, a subspace-projection technique is applied to the estimated coefficients to segregate the wall returns from the target coefficients [120]. Finally, a multitask linear model is developed which combines the single channel scene images with a composite scene image. All the scene images are recovered jointly using a multitask Bayesian approximation model.

6.3.1 Joint signal coefficient estimation

We introduce a joint Bayesian sparse model for signal coefficient estimation by exploiting both the signal sparsity and inter-signal correlations. In this approach, the intra-signal sparsity structure is enforced by imposing a sparseness prior on the signal coefficient vectors (model parameters). The inter-signal correlations are modeled by imposing a shared hyperparameter vector on the model parameters. The hyperparameters are used to control the common sparsity structure among the coefficient vectors. The model parameters and hyperparameters are estimated by maximizing the marginal likelihood and combining all measurement sets.

The signal model in Eqs. (6.8)–(6.9) represents for TWRI where a full data volume is collected using all M antenna locations and N frequencies at each channel. In compressed TWR sensing, only a subset of antennas are used for data collection, and at each selected antenna location, only a reduced set of frequencies are available for sensing the scene. Assume that only M_a ($M_a < M$) antenna locations are selected at each channel. At each selected antenna, only N_f ($N_f \ll N$) frequencies are used for illuminating the scene. Due to the constraint of compressive sampling, the frequencies are different from one antenna to another. The signal $\mathbf{z}_{l,i_a}^{wt}$ due to wall and target returns can be sparsely represented using a

dictionary $\mathbf{W} \in \mathbb{R}^{N \times J}$ containing J ($J \geq N$) basis functions or atoms,

$$\mathbf{z}_{l,i_a}^{wt} = \mathbf{W}\,\theta_{l,i_a}^{wt}, \tag{6.10}$$

where θ_{l,i_a}^{wt} is a vector of signal coefficients corresponding to the i_a-th antenna and the l-th channel.

In compressed multi-channel TWRI, instead of collecting the full N samples at each antenna, only a reduced set of N_f ($N_f \ll N$) measurements is acquired at the i_a-th antenna location. Let $\mathbf{y}_{l,i_a}$ be a vector containing the compressed frequency samples. The sparse sampling process can be mathematically expressed as

$$\mathbf{y}_{l,i_a} = \mathbf{\Phi}_{l,i_a}\,\mathbf{z}_{l,i_a}, \tag{6.11}$$

where $\mathbf{\Phi}_{l,i_a}$ is an $N_f \times N$ selection matrix containing a single unit value in each row and each column. Here, the selection matrix $\mathbf{\Phi}_{l,i_a}$ has a similar structure to an identity matrix, where each row contains only one non-zero element representing the selected frequency. From Eqs. (6.6) and (6.10), it follows that

$$\mathbf{y}_{l,i_a} = \mathbf{\Phi}_{l,i_a}\,\mathbf{W}\,\theta_{l,i_a}^{wt} + \mathbf{\Phi}_{l,i_a}\,\nu_{l,i_a} = \mathbf{D}_{l,i_a}\,\theta_{l,i_a}^{wt} + \epsilon_{l,i_a}, \tag{6.12}$$

where $\mathbf{D}_{l,i_a} = \mathbf{\Phi}_{l,i_a}\,\mathbf{W}$ and $\epsilon_{l,i_a} = \mathbf{\Phi}_{l,i_a}\,\nu_{l,i_a}$.

Given the measurement vector $\mathbf{y}_{l,i_a}$ and the dictionary $\mathbf{D}_{l,i_a}$, the coefficient vector θ_{l,i_a}^{wt} can be recovered using different approaches. In [91, 108], the vector θ_{l,i_a}^{wt} is recovered independently at each antenna. These methods, however, do not consider the correlations between antenna signals. In contrast, here we consider the multitask inverse problem (6.12) as a joint sparsity model, which assumes that all the coefficient vectors θ_{l,i_a}^{wt} have overlapping support. By exploiting the joint sparsity structure, the accuracy of signal reconstruction can be enhanced significantly. Several algorithms have been proposed which take the inter-signal correlations into account [123, 124, 125]. However, these methods assume that the

compressed measurements are obtained by using the same sensing matrix $\boldsymbol{\Phi}_{l,i_a}$. In the TWRI context, this assumption corresponds to a constraint of having the same frequency measurements for all antennas at all channels, and thus this is not possible in TWR sensing due to frequency interferences and radar jamming. To relax this constraint, we propose to use a joint Bayesian sparse recovery framework for jointly estimating all the coefficient vectors $\boldsymbol{\theta}_{l,i_a}^{wt}$.

In the sparse Bayesian model, the noise term in Eq. (6.12) is modeled as a complex Gaussian distribution, $p(\boldsymbol{\epsilon}_{l,i_a}) \sim \mathcal{CN}(\mathbf{0}, \beta\mathbf{I})$, with zero-mean and covariance $\beta\mathbf{I}$. The likelihood of $\boldsymbol{\theta}_{l,i_a}^{wt}$ is therefore a multivariate complex Gaussian function, which can be expressed as [52, 126]

$$p(\mathbf{y}_{l,i_a}|\boldsymbol{\theta}_{l,i_a}^{wt}, \beta) = (\pi/\beta)^{-N_f} \exp\left(-\beta \left\|\mathbf{y}_{l,i_a} - \mathbf{D}_{l,i_a}\boldsymbol{\theta}_{l,i_a}^{wt}\right\|_2^2\right). \tag{6.13}$$

The joint sparsity of the coefficient vectors is enforced using a shared complex Gaussian prior imposed on $\boldsymbol{\theta}_{l,i_a}^{wt}$:

$$p(\boldsymbol{\theta}_{l,i_a}^{wt}|\boldsymbol{\alpha}) = \prod_{j=1}^{J} \mathcal{CN}(\theta_{l,i_a}^{wt}(j) \mid 0, \alpha_j^{-1}), \tag{6.14}$$

$$= \prod_{j=1}^{J} \frac{1}{\pi\,\alpha_j} \exp\left(-\frac{|\theta_{l,i_a}^{wt}(j)|^2}{\alpha_j}\right) = \mathcal{CN}(\mathbf{0}, \mathbf{A}),$$

where $\mathbf{A} = \mathrm{diag}(\boldsymbol{\alpha})$ and $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_J]$ is the vector of hyperparameters. Note that the hyperparameter vector $\boldsymbol{\alpha}$ is used to capture the correlations among coefficient vectors $\boldsymbol{\theta}_{l,i_a}^{wt}$, controlling which elements of the vectors $\boldsymbol{\theta}_{l,i_a}^{wt}$ are non-zero or which basis functions of the dictionary $\mathbf{D}_{l,i_a}$ are selected.

Given the hyperparameters $\boldsymbol{\alpha}$ and β, by Bayes rule, the posterior of $\boldsymbol{\theta}_{l,i_a}^{wt}$ is a

multivariate complex Gaussian distribution:

$$p(\boldsymbol{\theta}_{l,i_a}^{wt}|\mathbf{y}_{l,i_a}, \boldsymbol{\alpha}, \beta) = \frac{p(\mathbf{y}_{l,i_a}|\boldsymbol{\theta}_{l,i_a}^{wt}, \beta)\, p(\boldsymbol{\theta}_{l,i_a}^{wt}|\boldsymbol{\alpha})}{p(\mathbf{y}_{l,i_a}|\boldsymbol{\alpha}, \beta)}, \tag{6.15}$$

$$= \frac{1}{(\pi)^{N_f}|\boldsymbol{\Sigma}_{l,i_a}|} \exp\left\{-(\boldsymbol{\theta}_{l,i_a}^{wt} - \boldsymbol{\mu}_{l,i_a})^H \boldsymbol{\Sigma}_{l,i_a}^{-1} (\boldsymbol{\theta}_{l,i_a}^{wt} - \boldsymbol{\mu}_{l,i_a})\right\}, \tag{6.16}$$

with mean and covariance given by

$$\boldsymbol{\mu}_{l,i_a} = \beta\, \boldsymbol{\Sigma}_{l,i_a}\, \mathbf{D}_{l,i_a}^H\, \mathbf{y}_{l,i_a}, \tag{6.17}$$

$$\boldsymbol{\Sigma}_{l,i_a} = (\beta\, \mathbf{D}_{l,i_a}^H\, \mathbf{D}_{l,i_a} + \mathbf{A})^{-1}. \tag{6.18}$$

The problem now becomes searching for the hyperparameters $\boldsymbol{\alpha}$ and β, which can be obtained by maximizing the logarithm of the marginal likelihood:

$$\mathcal{L}(\boldsymbol{\alpha}, \beta) = \sum_{a=1}^{M_a} \log p(\mathbf{y}_{l,i_a}|\boldsymbol{\alpha}, \beta), \tag{6.19}$$

$$= \sum_{a=1}^{M_a} \log \int p(\mathbf{y}_{l,i_a}|\boldsymbol{\theta}_{l,i_a}^{wt}, \beta)\, p(\boldsymbol{\theta}_{l,i_a}^{wt}|\boldsymbol{\alpha})\, d\boldsymbol{\theta}_{l,i_a}^{wt}, \tag{6.20}$$

$$= -\frac{1}{2} \sum_{a=1}^{M_a} \left[N_f \log 2\pi + \log|\mathbf{B}_{l,i_a}| + \mathbf{y}_{l,i_a}^H\, \mathbf{B}_{l,i_a}^{-1}\, \mathbf{y}_{l,i_a} \right], \tag{6.21}$$

with $\mathbf{B}_{l,i_a} = \beta^{-1}\mathbf{I} + \mathbf{D}_{l,i_a}\, \mathbf{A}^{-1}\, \mathbf{D}_{l,i_a}^H$.

An expectation-maximization algorithm is used to maximize $\mathcal{L}(\boldsymbol{\alpha}, \beta)$ with respect to the hyperparameters $\boldsymbol{\alpha}$ and β. This algorithm is an efficient estimation procedure that iterates among expectation (E) and maximization (M) steps. In the E-step, the iterative algorithm estimates the posterior mean (6.17) and covariance (6.18). For M-step, by taking the derivatives of the cost function $\mathcal{L}(\boldsymbol{\alpha}, \beta)$ with respect to $\boldsymbol{\alpha}$ and β and setting the results to zero, we derive the following update

rules:

$$\hat{\alpha}_j = \frac{M_a - \alpha_j \sum_{a=1}^{M_a} \Sigma_{l,i_a}(j,j)}{\sum_{a=1}^{M_a} |\mu_{l,i_a}(j)|^2}, \tag{6.22}$$

$$\hat{\beta} = \frac{\sum_{a=1}^{M_a} \left[N_f - J + \sum_{j=1}^{J} \alpha_j \, \Sigma_{l,i_a}(j,j) \right]}{\sum_{a=1}^{M_a} \|\mathbf{y}_{l,i_a} - \mathbf{D}_{l,i_a} \, \boldsymbol{\mu}_{l,i_a}\|_2^2}, \tag{6.23}$$

where $\mu_{l,i_a}(j)$ is the j-th component of $\boldsymbol{\mu}_{l,i_a}$, and $\Sigma_{l,i_a}(j,j)$ is the j-th diagonal element of $\boldsymbol{\Sigma}_{l,i_a}$. The algorithm iterates between (6.17)-(6.18) and (6.22)-(6.23) until convergence is achieved or a stopping criterion is satisfied [52, 63]. Once the hyperparameters $\hat{\boldsymbol{\alpha}}$ and $\hat{\beta}$ are estimated, the signal coefficient vector $\hat{\boldsymbol{\theta}}_{l,i_a}^{wt}$ is given by the mean of the posterior in (6.17):

$$\hat{\boldsymbol{\theta}}_{l,i_a}^{wt} = \boldsymbol{\mu}_{l,i_a}|_{\alpha=\hat{\alpha}, \, \beta=\hat{\beta}} = (\beta \, \boldsymbol{\Sigma}_{l,i_a} \, \mathbf{D}_{l,i_a}^{H} \, \mathbf{y}_{l,i_a})|_{\alpha=\hat{\alpha}, \, \beta=\hat{\beta}}. \tag{6.24}$$

Here, it is important to note that all the compressed measurement sets $\{\mathbf{y}_{l,i_a}\}_{a=1}^{M_a}$ are employed for the estimation of the hyperparameters, as shown by the summation of the conditional distributions in (6.19). Hence, the correlations between the antenna signals are exploited through the learning of the hyperparameters. Once the signal coefficient vectors $\hat{\boldsymbol{\theta}}_{l,i_a}^{wt}$ are obtained, the scene image reconstruction can proceed. However, the estimated coefficient vectors $\hat{\boldsymbol{\theta}}_{l,i_a}^{wt}$ contain the wall-reverberation components, which usually dominate the target signal and hinder the visibility of stationary targets in the image. Therefore, before image reconstruction, the wall reverberations need to be identified and suppressed.

6.3.2 Wall coefficient mitigation

Usually, wall mitigation techniques are applied to the radar signals [96, 119, 120], which can be recovered from the estimated coefficients $\hat{\boldsymbol{\theta}}_{l,i_a}^{wt}$ using Eq. (6.10). In the proposed approach, however, we apply a subspace-projection method directly to the estimated coefficients to segregate the wall contributions from the target returns. Let $\hat{\boldsymbol{\Theta}}^{wt}$ denote a matrix comprising in its columns the antenna coefficients

$\hat{\theta}_{l,i_a}^{wt}$ obtained from all channels. Using singular value decomposition, the matrix $\hat{\Theta}^{wt}$ can be expressed as

$$\hat{\Theta}^{wt} = \mathbf{U} \, \mathbf{\Sigma} \, \mathbf{V}^H, \tag{6.25}$$

where $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_J]$ and $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_{M_a L}]$ are unitary matrices containing the left and right singular vectors, respectively, and $\mathbf{\Sigma}$ is a matrix containing the singular values along the main diagonal arranged in descending order.

In TWRI, the wall returns are relatively stronger than the target reflections. Hence, the wall contributions are captured by the first few singular vectors associated with the dominant singular values. The wall subspace can be defined as

$$\mathbf{P}_w = \sum_{i \in \mathcal{W}} \mathbf{u}_i \, \mathbf{v}_i^H, \tag{6.26}$$

where $\mathcal{W}$ denotes the index set of the singular vectors spanning the wall subspace determined by a classification technique [119]. To suppress the wall coefficients, the matrix $\hat{\Theta}^{wt}$ is projected onto a subspace orthogonal to the wall subspace:

$$\tilde{\Theta}^t = (\mathbf{I} - \mathbf{P}_w \mathbf{P}_w^H) \, \hat{\Theta}^{wt}, \tag{6.27}$$

where $\mathbf{I}$ denotes the identity matrix. Now the wall-clutter free coefficients $\tilde{\Theta}^t$, which contain in its columns the target coefficient vectors $\tilde{\theta}_{l,i_a}^t$, can be used for image reconstruction.

6.3.3 Joint Bayesian sparse scene reconstruction

The scene can be formed by first reconstructing the radar signal from the target coefficients $\tilde{\theta}_{l,i_a}^t$ using (6.10), and then applying DS beamforming [91], or ℓ_1 minimization [108, 122]. However, these methods are designed for single-view image formation, which ignore the inter-channel dependencies. Here, we formulate a multitask linear model that maps the clutter-free coefficients to the corresponding images of the scene and incorporates a composite coefficient vector representing the fused image of the scene.

Let $\tilde{\boldsymbol{\theta}}_{l,i_a}^t$, for $a = 0, 1, \ldots, M_a - 1$ and $l = 1, 2, \ldots, L$ be the remaining signal coefficients after removing those associated with the wall returns. Using Eqs. (6.7) and (6.10), we can relate the target coefficients to the l-th channel target image of the scene $\mathbf{s}_l^t$ as

$$\tilde{\boldsymbol{\theta}}_{l,i_a}^t = \mathbf{W}^\dagger \, \boldsymbol{\Psi}_{l,i_a} \, \mathbf{s}_l^t + \boldsymbol{\xi}_{l,i_a}, \tag{6.28}$$

where $\dagger$ denotes the pseudo-inverse operator and $\boldsymbol{\xi}_{l,i_a}$ is the residual noise. The target coefficient vectors $\tilde{\boldsymbol{\theta}}_{l,i_a}^t$ obtained from all selected antenna locations for the l-th channel can be concatenated to form a composite measurement vector, which can be used to reconstruct the corresponding image of the scene of the l-th channel. However, this approach yields a very high-dimensional measurement vector of size $M_a J \times 1$ and a huge dictionary of size $M_a J \times Q$, which leads to inefficient image reconstruction. We therefore apply a principal component analysis technique to compress the measurement vector and the dictionary and obtain a compact linear model.

Let $\mathbf{C}$ denote the covariance matrix of the new measurements given in Eq. (6.28). Using the eigendecomposition, we can write

$$\mathbf{C} = \mathbf{G} \, \boldsymbol{\Lambda} \, \mathbf{G}^H, \tag{6.29}$$

where $\mathbf{G} = [\mathbf{g}_0, \ldots, \mathbf{g}_{J-1}]$ is a matrix of eigenvectors and $\boldsymbol{\Lambda}$ is a diagonal matrix of eigenvalues, arranged in descending order. We define a projection matrix $\mathbf{P} = [\mathbf{g}_0, \ldots, \mathbf{g}_{N_p-1}]$ consisting of the first N_p $(N_p \ll J)$ eigenvectors; N_p is determined by using an information theoretic criterion, such as Akaike Information Criterion (AIC) or Minimum Description Length (MDL) [128]. Applying the projection matrix $\mathbf{P}^H$ to Eq. (6.28) yields the compressed signal coefficient vector $\tilde{\mathbf{y}}_{i_a,l} \in \mathbb{C}^{N_p}$,

$$\tilde{\mathbf{y}}_{l,i_a} = \mathbf{P}^H \tilde{\boldsymbol{\theta}}_{l,i_a}^t = \mathbf{P}^H \mathbf{W}^\dagger \, \boldsymbol{\Psi}_{l,i_a} \, \mathbf{s}_l^t + \mathbf{P}^H \boldsymbol{\xi}_{l,i_a} = \tilde{\boldsymbol{\Psi}}_{l,i_a} \, \mathbf{s}_l^t + \tilde{\boldsymbol{\xi}}_{l,i_a}, \tag{6.30}$$

where $\tilde{\boldsymbol{\Psi}}_{l,i_a} = \mathbf{P}^H \mathbf{W}^\dagger \, \boldsymbol{\Psi}_{l,i_a}$ and $\tilde{\boldsymbol{\xi}}_{l,i_a} = \mathbf{P}^H \boldsymbol{\xi}_{l,i_a}$. Now, we stack all the compressed

signal coefficient vectors belonging to the l-th channel into a composite vector $\tilde{\mathbf{y}}_l = [\tilde{\mathbf{y}}_{l,0}^T, \dots, \tilde{\mathbf{y}}_{l,M_a-1}^T]^T \in \mathbb{C}^{M_a N_p \times 1}$, arrange the individual compressed dictionaries into a composite dictionary $\tilde{\boldsymbol{\Psi}}_l = [\tilde{\boldsymbol{\Psi}}_{l,0}^T, \dots, \tilde{\boldsymbol{\Psi}}_{l,M_a-1}^T]^T \in \mathbb{C}^{M_a N_p \times Q}$, and noise component $\tilde{\boldsymbol{\xi}}_l = [\tilde{\boldsymbol{\xi}}_{l,0}^T, \dots, \tilde{\boldsymbol{\xi}}_{l,M_a-1}^T]^T$. We can obtain the following compact linear model

$$\tilde{\mathbf{y}}_l = \tilde{\boldsymbol{\Psi}}_l \, \mathbf{s}_l^t + \tilde{\boldsymbol{\xi}}_l. \tag{6.31}$$

In the linear model (6.31), the scene vectors $\mathbf{s}_l^t$ have a common sparsity support since they represent images of the same scene targets. Note that for multi-view sensing, we need to adopt a pixel scanning scheme in which all the vectors $\mathbf{s}_l^t$ have the same sparsity support [138] since the imaging coordinates are different between viewing angles. Now, the scene images $\mathbf{s}_l^t$ can be jointly reconstructed using the multitask Bayesian CS technique that considers the inter-channel correlations. Furthermore, because the vectors $\mathbf{s}_l^t$ represent images of the same scene, a composite image of the scene can be obtained by using image fusion techniques after each single-channel image has been reconstructed and aligned. Here instead we propose to first combine the coefficient vectors from different channels, then perform fusion using joint Bayesian sparse learning.

We formulate a composite coefficient vector $\bar{\mathbf{y}}_l$ as a linear combination of the coefficient vectors of different channels: $\bar{\mathbf{y}} = \sum_{l=1}^{L} w_l \, \tilde{\mathbf{y}}_l$, where w_l's are positive weights satisfying $\sum_{l=1}^{L} w_l = 1$, computed based on mutual information (MI) [116]. Similarly, we can form a composite dictionary as $\bar{\boldsymbol{\Psi}} = \sum_{l=1}^{L} w_l \tilde{\boldsymbol{\Psi}}_l$. A linear imaging model relating the composite coefficient vector $\bar{\mathbf{y}}$ to the fused image of the scene $\bar{\mathbf{s}}^t$ can be formulated as $\bar{\mathbf{y}} = \bar{\boldsymbol{\Psi}} \, \bar{\mathbf{s}}^t$. By combining the composite linear model with (6.31), we obtain an overall multitask model for the multi-channel TWRI problem:

$$\tilde{\mathbf{y}}_l = \tilde{\boldsymbol{\Psi}}_l \, \mathbf{s}_l^t + \tilde{\boldsymbol{\xi}}_l, \; l = 1, \dots, (L+1), \tag{6.32}$$

where $\tilde{\mathbf{y}}_{L+1} = \bar{\mathbf{y}}$, $\tilde{\boldsymbol{\Psi}}_{L+1} = \bar{\boldsymbol{\Psi}}$, and $\tilde{\mathbf{s}}_{L+1} = \bar{\mathbf{s}}^t$. The solution of (6.32) yields $(L+1)$

images corresponding to the L individual channels plus a composite image of the scene. This multitask problem can be solved efficiently using the joint Bayesian sparse model, taking the inter-channel correlations into account.

In the sparse Bayesian framework, the noise vector in (6.32) is modeled as $\tilde{\xi}_l \sim \mathcal{CN}(\mathbf{0}, \tilde{\beta}\mathbf{I})$, with zero-mean and covariance $\tilde{\beta}\mathbf{I}$. Then, the likelihood for $\mathbf{s}_l^t$ is a multivariate complex Gaussian distribution:

$$p(\tilde{\mathbf{y}}_l|\mathbf{s}_l^t, \tilde{\beta}) = (\pi\tilde{\beta})^{-M_a N_p} \exp\left(-\tilde{\beta}\|\tilde{\mathbf{y}}_l - \tilde{\mathbf{\Psi}}_l\,\mathbf{s}_l^t\|_2^2\right). \tag{6.33}$$

The inter-channel correlation structure is enforced using a shared sparseness prior:

$$p(\mathbf{s}_l^t|\tilde{\boldsymbol{\alpha}}) = \prod_{q=1}^{Q} \mathcal{CN}(s_l^t(q)\,|\,0, \tilde{\alpha}_q^{-1}), \tag{6.34}$$

$$= \prod_{q=1}^{Q} \frac{1}{\pi\tilde{\alpha}_q} \exp\left(-\frac{|s_l^t(q)|^2}{\tilde{\alpha}_q}\right) = \mathcal{CN}(0|\tilde{\mathbf{A}}), \tag{6.35}$$

where $\tilde{\mathbf{A}} = \mathrm{diag}(\tilde{\boldsymbol{\alpha}})$, and $\tilde{\boldsymbol{\alpha}} = [\tilde{\alpha}_1, \ldots, \tilde{\alpha}_Q]$ is the vector of hyperparameters.

Given the hyperparameters $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\beta}$, by applying the Bayes rule, the posterior for $\mathbf{s}_l^t$ is a multivariate complex Gaussian distribution:

$$p(\mathbf{s}_l^t|\tilde{\mathbf{y}}_l, \tilde{\boldsymbol{\alpha}}, \tilde{\beta}) = \frac{p(\tilde{\mathbf{y}}_l|\mathbf{s}_l^t, \tilde{\beta})\,p(\mathbf{s}_l^t|\tilde{\boldsymbol{\alpha}})}{p(\tilde{\mathbf{y}}_l|\tilde{\boldsymbol{\alpha}}, \tilde{\beta})}, \tag{6.36}$$

$$= \frac{1}{(\pi)^{M_a N_p}|\tilde{\mathbf{\Sigma}}_l|} \exp\left\{-(\mathbf{s}_l^t - \tilde{\boldsymbol{\mu}}_l)^H \tilde{\mathbf{\Sigma}}_l^{-1} (\mathbf{s}_l^t - \tilde{\boldsymbol{\mu}}_l)\right\}, \tag{6.37}$$

with the mean and covariance given by

$$\tilde{\boldsymbol{\mu}}_l = \tilde{\beta}\,\tilde{\mathbf{\Sigma}}_l\,\tilde{\mathbf{\Psi}}_l^H\,\tilde{\mathbf{y}}_l, \tag{6.38}$$

$$\tilde{\mathbf{\Sigma}}_l = (\tilde{\beta}\,\tilde{\mathbf{\Psi}}_l^H\,\tilde{\mathbf{\Psi}}_l + \tilde{\mathbf{A}})^{-1}. \tag{6.39}$$

The learning problem now is to search for the hyperparameters $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\beta}$. These hyperparameters are estimated by maximizing the marginal likelihood, or

equivalently its logarithm:

$$\mathcal{L}(\tilde{\boldsymbol{\alpha}}, \tilde{\beta}) = \sum_{l=1}^{L+1} \log p(\tilde{\mathbf{y}}_l | \tilde{\boldsymbol{\alpha}}, \tilde{\beta}), \tag{6.40}$$

$$= \sum_{l=1}^{L+1} \log \int p(\tilde{\mathbf{y}}_l | \mathbf{s}_l^t, \tilde{\beta}) \, p(\mathbf{s}_l^t | \tilde{\boldsymbol{\alpha}}) \, d\mathbf{s}_l^t, \tag{6.41}$$

$$= -\frac{1}{2} \sum_{l=1}^{L+1} \left[M_a N_p \log 2\pi + \log |\tilde{\mathbf{B}}_l| + \tilde{\mathbf{y}}_l^H \, \tilde{\mathbf{B}}_l^{-1} \, \tilde{\mathbf{y}}_l \right], \tag{6.42}$$

where $\tilde{\mathbf{B}}_l = \tilde{\beta}^{-1}\mathbf{I} + \tilde{\boldsymbol{\Psi}}_l \, \tilde{\mathbf{A}}^{-1} \, \tilde{\boldsymbol{\Psi}}_l^H$.

The expectation-maximization algorithm is used to maximize $\mathcal{L}(\tilde{\boldsymbol{\alpha}}, \tilde{\beta})$ with respect to $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\beta}$. For E-step, it computes the posterior mean (6.38) and covariance (6.39). For M-step, it takes the derivatives of the cost function $\mathcal{L}(\tilde{\boldsymbol{\alpha}}, \tilde{\beta})$ with respect to $\tilde{\boldsymbol{\alpha}}$ and $\tilde{\beta}$, leading to the following update rules:

$$\bar{\alpha}_q = \frac{L + 1 - \tilde{\alpha}_q \sum_{l=1}^{L+1} \tilde{\Sigma}_l(q, q)}{\sum_{l=1}^{L+1} |\tilde{\mu}_l(q)|^2}, \tag{6.43}$$

$$\bar{\beta} = \frac{\sum_{l=1}^{L+1} \left[M_a N_p - Q + \sum_{q=1}^{Q} \tilde{\alpha}_q \tilde{\Sigma}_l(q, q) \right]}{\sum_{l=1}^{L+1} \left\| \mathbf{y}_l - \tilde{\boldsymbol{\Psi}}_l \, \tilde{\boldsymbol{\mu}}_l \right\|_2^2}, \tag{6.44}$$

where $\tilde{\mu}_l(q)$ is the q-th component of $\tilde{\boldsymbol{\mu}}_l$, and $\tilde{\Sigma}_l(q, q)$ is the q-th diagonal element of $\tilde{\boldsymbol{\Sigma}}_l$. The algorithm iterates between (6.38)-(6.39) and (6.43)-(6.44) until convergence. Once the hyperparameters $\bar{\boldsymbol{\alpha}}$ and $\bar{\beta}$ are estimated, the scene image $\hat{\mathbf{s}}_l^t$ is given by the mean of the posterior in (6.38):

$$\hat{\mathbf{s}}_l^t = \tilde{\boldsymbol{\mu}}_l|_{\tilde{\boldsymbol{\alpha}}=\bar{\boldsymbol{\alpha}}, \, \tilde{\beta}=\bar{\beta}} = (\tilde{\beta} \, \tilde{\boldsymbol{\Sigma}}_l \, \tilde{\boldsymbol{\Psi}}_l^H \, \tilde{\mathbf{y}}_l)|_{\tilde{\boldsymbol{\alpha}}=\bar{\boldsymbol{\alpha}}, \, \tilde{\beta}=\bar{\beta}}. \tag{6.45}$$

Here, it is important to note that all the compressed coefficient sets $\tilde{\mathbf{y}}_l$, for $l = 1, \ldots, L + 1$, are combined to estimate the hyperparameters, as shown by the summation of the conditional distributions in Eqs. (6.40)–(6.42). Hence, the correlations between channels are exploited through learning of the hyperparameters and estimating of the posterior mean and covariance.

6.4 Experimental results and analysis

In this section, we present experimental results for multi-channel TWRI. Both multi-view and multi-polarization radar signals are used to validate the proposed approach. The performance of the proposed approach is compared with conventional single-channel imaging model where the image at each channel is formed separately, followed by fusion in the image domain to generate a fused image of the scene. Before presenting the imaging results, we give the experimental setup for simulations.

6.4.1 Experimental setup

To represent the radar signals sparsely, see Eq. (6.10), in all experiments, the dictionary $\mathbf{W}$ is constructed using modulated discrete prolate spheroidal sequences (DPSS) since the DPSS dictionary is overcomplete and can represent bandpass radar signals more compactly in the range profile than does the Fourier basis [122]. The formulation for the dictionary $\mathbf{W}$ is given in Section 5.5.

To quantify the performances of different imaging models, three common measures are used: the reconstruction error (RE), the target-to-clutter ratio (TCR), and the receiver operating characteristic (ROC) curve. The reconstruction error is calculated as

$$\text{RE} = \|\mathbf{z} - \hat{\mathbf{z}}\|_2 / \|\mathbf{z}\|_2, \tag{6.46}$$

where $\hat{\mathbf{z}}$ and $\mathbf{z}$ are the reconstructed and true signals, respectively. The target-to-clutter ratio (in dB) is computed as

$$\text{TCR} = 10 \log_{10}(P_{\text{target}}/P_{\text{clutter}}), \tag{6.47}$$

where P_{target} and P_{clutter} are the average power in the target and clutter regions, respectively. The ground-truth target regions are selected manually in the vicinity of the true targets, and the clutter region is the remainder of the reconstructed

image. The receiver operating characteristic (ROC) curve is used to measure the probability of target detection for a given false alarm rate. The probability of detection, or detection rate, denotes the percentage of pixels in target regions that are correctly detected. By contrast, the probability of false alarm, or false alarm rate, is the percentage of pixels in the clutter region that are incorrectly detected as targets.

6.4.2 Results of multi-view data

In this section, the proposed approach is evaluated using multi-view radar signals. The radar signals were generated by electromagnetic simulations using XFDTD, a full-wave EM simulator developed based on the Finite Difference Time Domain method [139]. We build a computer model of an enclosed structure scene containing three dihedral targets. Dihedral 1 and Dihedral 2 have orientation angle facing the front wall, whereas Dihedral 3 has its corner oriented towards the side wall. The scene behind a concrete wall is illuminated from two different aspect angles: 0° view (through the front wall) and 90° view (through the side wall of the enclosed structure), see Fig. 6.1. At each view, the transceiver is placed at 51 positions parallel to the wall at a standoff distance of 1 m, to synthesize an array aperture of length 1.2 m. The transmitted frequency range is 1 GHz, centered at 2.5 GHz, with a step frequency of 3 MHz (i.e. 334 frequency bins).

In the first experiment, we used half of the antennas (M_a = 25 out of 51 antennas) and varied the selected frequencies from 10% to 40% (N_f = 33 to 134 out of 334 frequencies) of the full frequencies at each view. For each set of measurements, the signal coefficients were recovered by the single-signal CS model and joint Bayesian CS model. The reconstruction errors were recorded for 100 trials.

Figure 6.2 shows the reconstruction errors for both models. Compared to the single-signal CS model, the proposed joint Bayesian CS model produces a considerably lower reconstruction error, especially when the number of measurements

115
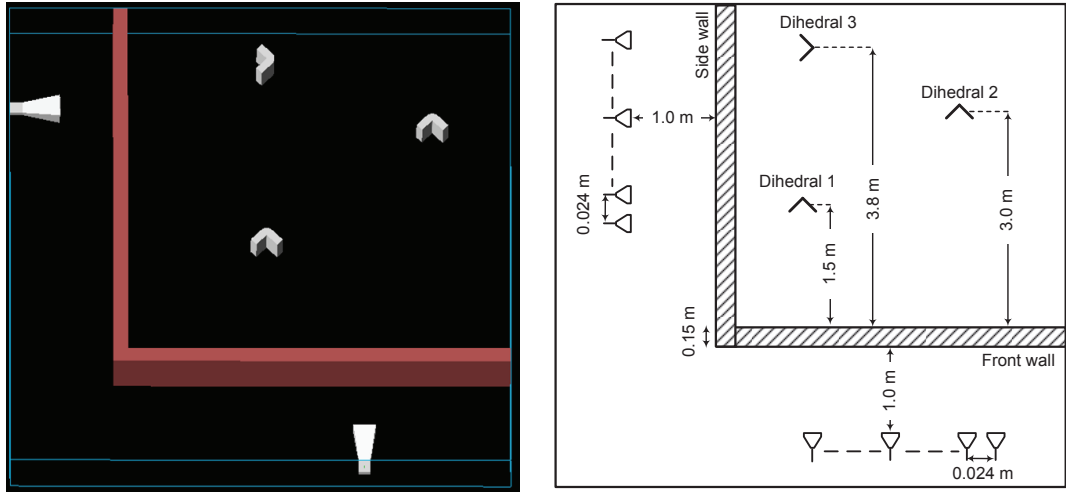
Figure 6.1: Multi-view TWRI data acquisition for an enclosed structure target scene. Left: the scene is imaged through a homogeneous concrete wall from two vantage points; Right: the schematic diagram of the scene.

is very low. At the ratio of 10% frequencies, the average reconstruction errors by the proposed joint Bayesian CS model and the conventional single-signal CS model are 0.05 and 0.4, respectively (an 8-fold improvement of reconstruction accuracy). Moreover, to obtain the same reconstruction accuracy, the proposed approach requires far fewer measurements than does the single-signal CS model. For example, to obtain an average RE = 0.05, the joint Bayesian sparse approach requires only 10% of the frequency measurements, whereas the single-signal CS model uses 25%. The reconstruction superiority of the proposed joint Bayesian model is because it exploits the signal sparsity and the correlations among the signals, whereas the conventional single-signal CS model considers only the signal sparsity.

In the second experiment, the signals recovered using 20% of the total measurements are used for scene reconstruction, after wall clutter mitigation. Figure 6.3 shows the images formed by different methods using the signals recovered with the single-signal CS model. Figures 6.3(a)–(b) present the front-view and side-view images formed using DS beamforming, averaged over 100 trials; these images contain heavy clutter and very weak targets. Figure 6.3(c) presents the composite image obtained from the images of Figs. 6.3(a) and (b) using MI-based
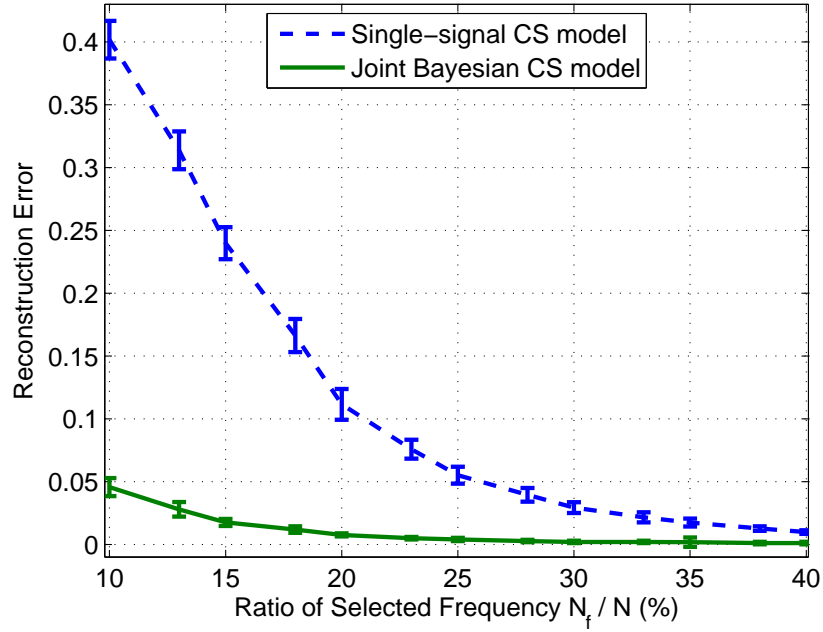
Figure 6.2: Reconstruction errors for signal coefficient estimation using the single-signal CS model and the proposed joint Bayesian CS model. The results are averaged over 100 trials. The error bars represent plus/minus one standard deviation.

fusion. In this case, the targets are not detectable, even in the fused image. Figures 6.3(d)–(f) illustrate the images obtained separately from the front view, side view, and their fused image, respectively, using the conventional CS reconstruction. Again, it is impossible to localize all the targets.

Figure 6.4 shows the scene images formed by different imaging methods using the signals jointly recovered by the proposed Bayesian sparse model. Figures 6.4(a)–(c) present, respectively, the front-view, side-view, and composite images reconstructed using standard CS model. Note that the composite image is formed by combining the measurements from the two views, but these images are recovered separately using the existing single-view CS method. The single-view images in Figs. 6.4(a) and (b) do not provide a complete picture of the scene content: Dihedrals 1 and 2 are weak in the side view, and Dihedral 3 is weak in the frontal view. Moreover, we can observe the outliers in the composite image shown in Fig. 6.4(c). By contrast, Figs. 6.4(d)–(f) show the images formed by the joint Bayesian sparse model; they contain much less clutter and reveal all
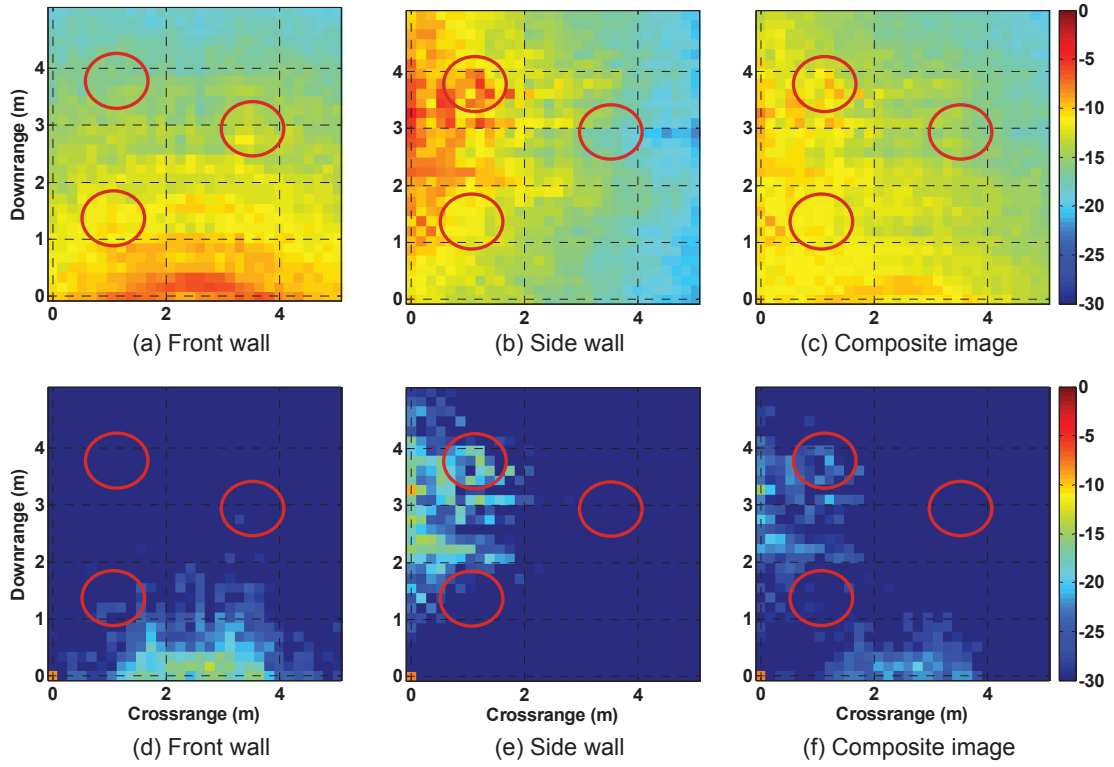
Figure 6.3: Images reconstructed from signals recovered by the single-signal CS model: (a)–(c) images reconstructed using the DS beamforming, (d)–(f) images reconstructed separately using the standard CS model.

the targets. The TCRs of the composite scene images formed by the conventional CS and the proposed Bayesian approach are 15.02 dB and 34.68 dB, respectively.

Figure 6.5 illustrates the ROC curves of the different imaging models. This figure shows that by reconstructing multiple images jointly and exploiting the inter-view dependencies, target detection is significantly enhanced, compared with the methods that form the images individually at each view.

6.4.3 Results of multi-polarization data

We also use radar signals acquired from a single vantage point using co-horizontal (HH) and co-vertical (VV) polarizations for evaluating the proposed approach. The horizontal and vertical polarization data sets were collected from a dihedral scene illustrated in Fig. 5.8. It contained a 0.4 m high and 0.3 m wide dihedral, placed on a turntable made of two 1.2 m x 2.4 m sheets of 0.013 m thick plywood. For each polarization setting, the scene is illuminated by a stepped-frequency
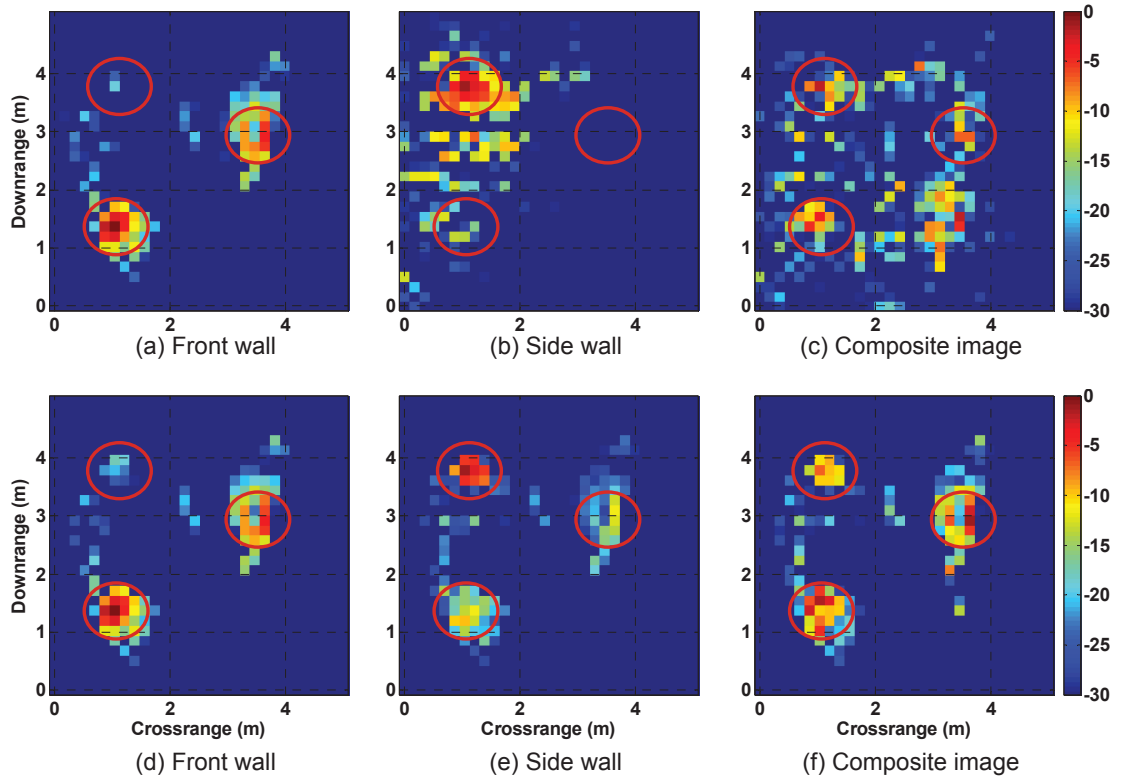
Figure 6.4: The scene reconstructed using signals jointly recovered by the proposed Bayesian model: (a)–(c) images reconstructed separately using the standard CS model, (d)–(f) images formed using the joint Bayesian sparse reconstruction.
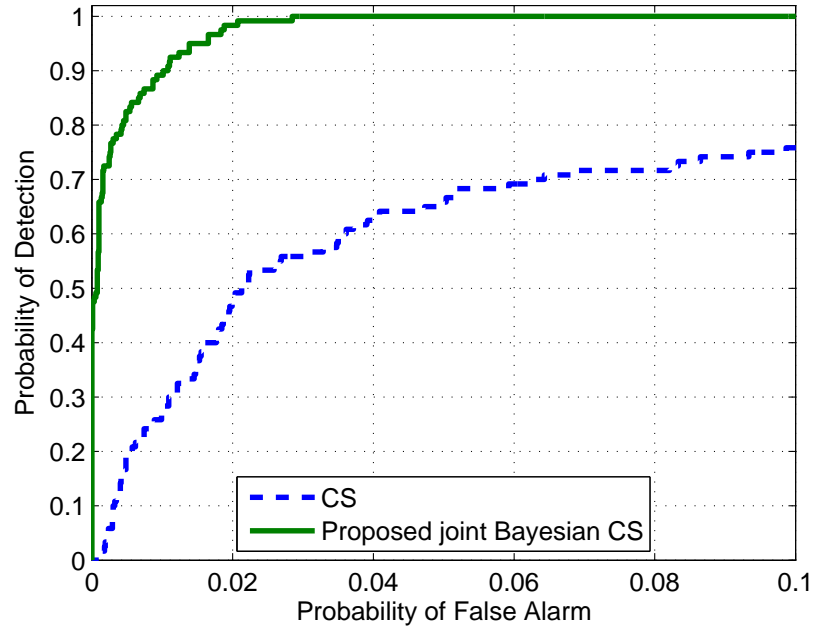
Figure 6.5: ROC curves of target detection on composite images formed by the standard CS model and the proposed joint Bayesian CS model.

synthetic aperture radar (SAR) system. The synthetic linear aperture consisted of 57 uniformly spaced elements, with inter-element spacing of 0.022 m. The antenna aperture was located at a height of 1.22 m above the floor and a standoff distance of 1.016 m away from a 0.143 m-thick solid concrete wall. The stepped-frequency signal comprised 801 frequencies, ranged between 0.7 and 3.1 GHz, with 3 MHz frequency step. The imaged scene extended from [0, 4] m in downrange and [−2, 2] m in crossrange. The pixel size was equal to the Rayleigh resolution of the radar, which resulted in a scene image of size 64×25 pixels.
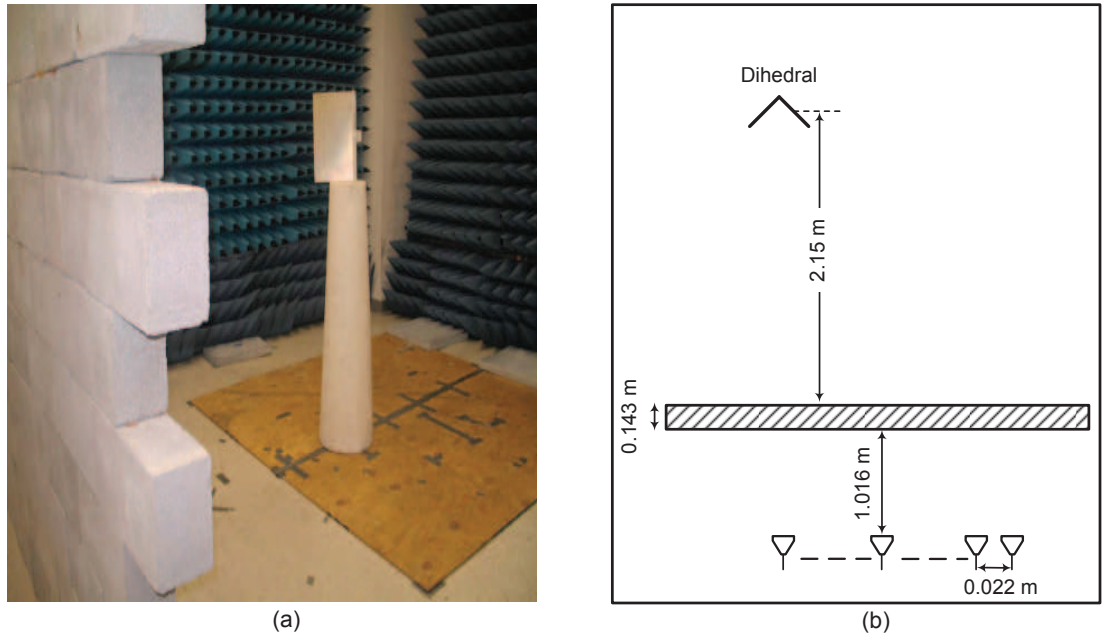
Figure 6.6: Multi-polarization TWRI data acquisition: (a) a photo of the scene; (b) a top-view of the behind-the-wall scene.

In the first experiment, the performance of the proposed approach is evaluated by using several compressed measurement sets. We selected $M_a = 6$ out of 57 antennas (10% of the total antennas) at each channel and varied the frequencies from 10% to 25% of the full frequencies ($N_f = 80, \ldots, 200$ out of 801). Using the compressive data sets, the signal coefficients were reconstructed by the single-signal CS model and the proposed joint Bayesian CS model. The reconstruction errors were recorded for 100 trials.

Figure 6.7 shows the average reconstruction error by different CS models as a

function of the ratio of selected frequencies. It is observed that the proposed joint Bayesian CS model reconstructs the antenna signals with substantially lower errors than does the single-signal CS model. This observation applies to all the sampling rates, and is more noticeable at low sampling rates. At the ratio $N_f/N = 10\%$, the average reconstruction errors by the single-signal CS model and the proposed joint Bayesian model are 0.135 and 0.039, respectively. Furthermore, the reconstruction by the proposed approach is much more stable than the single-signal CS model. The standard deviations of the reconstruction errors by the single-signal CS model and the proposed joint Bayesian model are, respectively, 0.011 and 0.002 at $N_f/N = 10\%$.
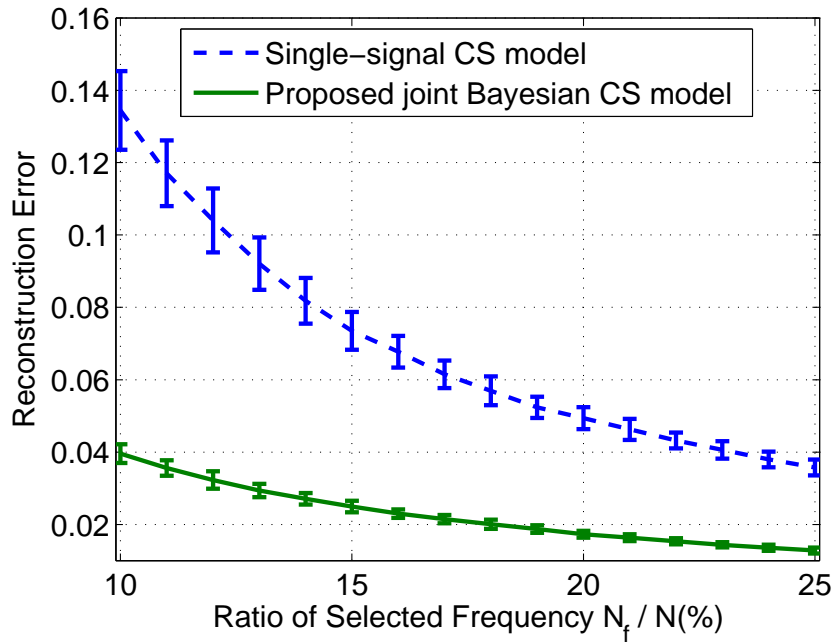
Figure 6.7: Reconstruction errors of multi-channel dihedral scene signals by different imaging models as a function of the percentage of selected frequencies. The results are averaged over 100 trials. The error bars represent plus/minus one standard deviation.

In the second experiment, the signal coefficients reconstructed from $N_f/N = 10\%$ of frequencies (collectively representing for just 1% of the total data volume at each polarimetric channel) are used for image formation, after wall-clutter mitigation. Figure 6.8 shows the polarimetric images formed by different methods

using the signals recovered with the single-signal CS. Figures 6.8(a)–(b) present the HH and VV images formed using DS beamforming, averaged over 100 trials. In these images, the target is shadowed due to heavy clutter. Figure 6.8(c) presents the composite image obtained from the images of Figs. 6.8(a) and (b) using MI-based fusion. Clearly, it is impossible to detect the dihedral target, even in the fused image. Figures 6.8(d)–(f) illustrate the images obtained separately from the HH channel, VV channel, and their composite image, respectively, using the conventional CS reconstruction. Again, in these formed images, it is impossible to localize the target.
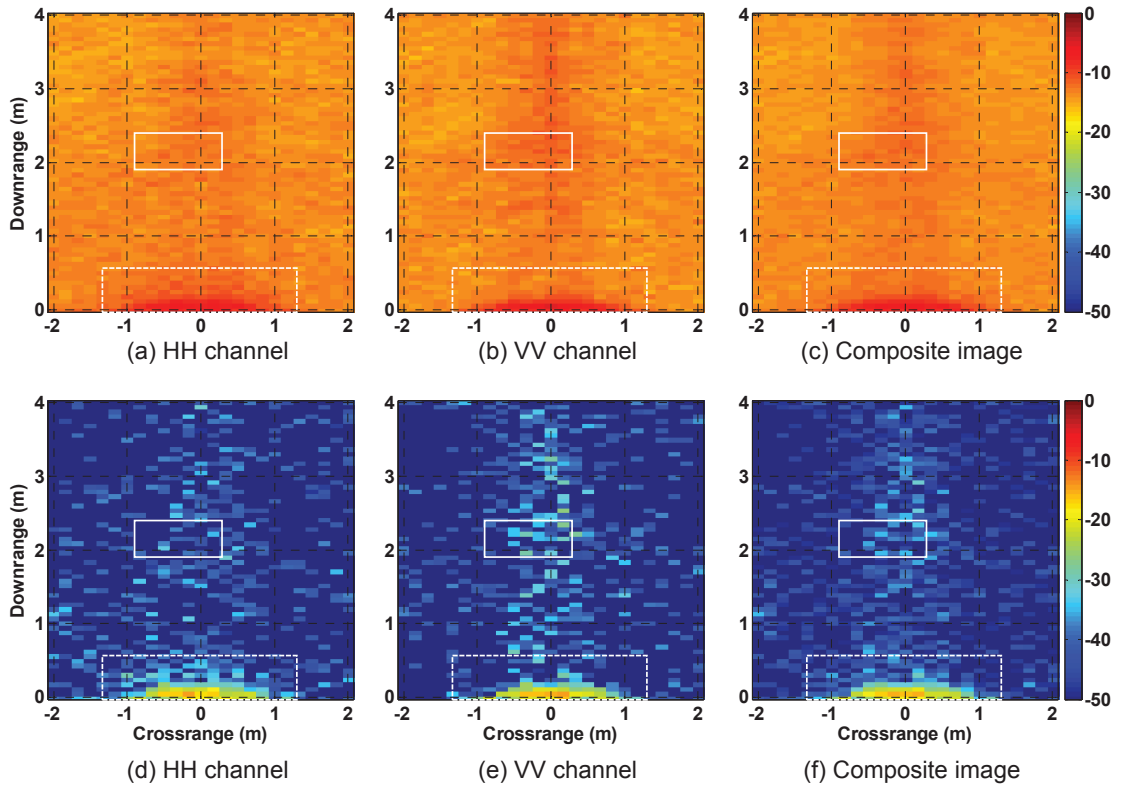
Figure 6.8: Polarimetric images reconstructed from signals recovered by single-signal CS model: (a)–(c) images reconstructed using the DS beamforming, (d)–(f) images reconstructed separately using the standard CS model.

Figure 6.9 shows polarimetric images formed by different imaging methods using the signals jointly recovered by the proposed Bayesian sparse model. Figures 6.9(a)–(c) present, respectively, the HH, VV, and composite images reconstructed using standard CS model. Note that the composite image is formed by
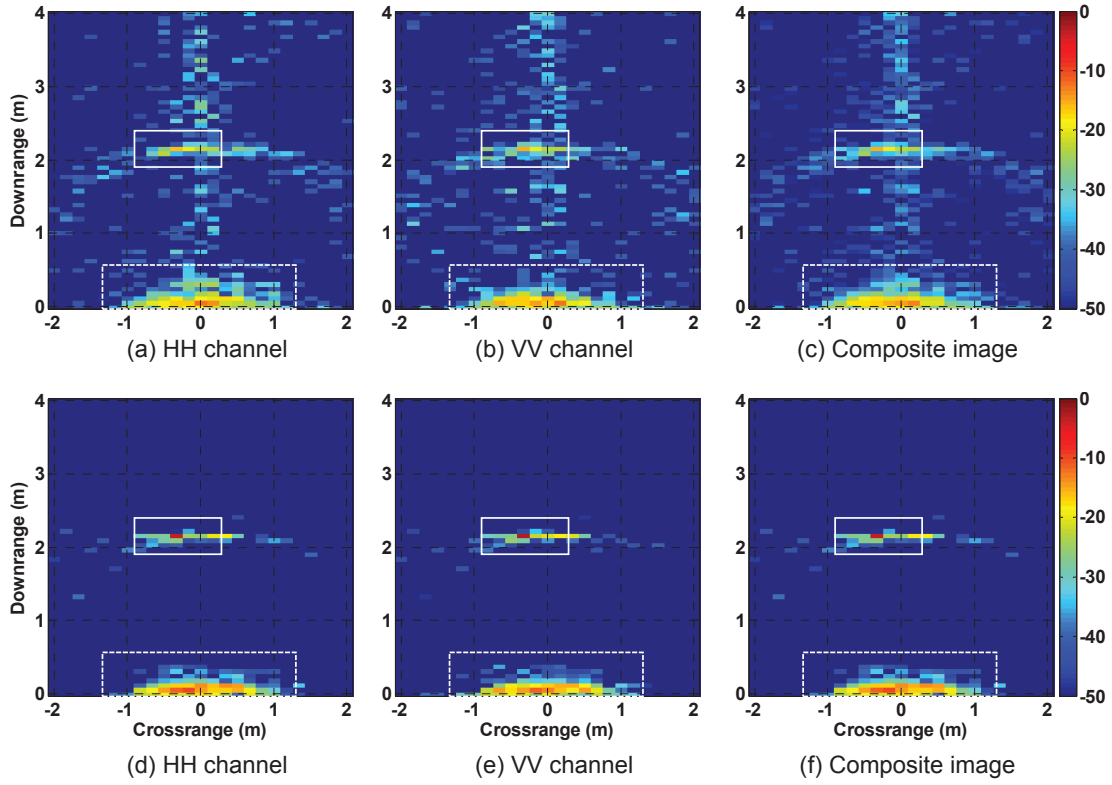
Figure 6.9: Polarimetric images reconstructed using signals jointly recovered by the proposed Bayesian model: (a)–(c) images reconstructed separately using the standard CS model, (d)–(f) images formed using the joint Bayesian sparse reconstruction.

combining the measurements from the two polarimetric channels, but these images are recovered separately using the existing single-channel CS method. The single-channel images in Figs. 6.9(a) and (b) do not localize the dihedral target well: the intensities in the target region are as weak as that of the outliers. The same effect can be observed in the composite image shown in Fig. 6.9(c). By contrast, Figs. 6.9(d)–(f) show the images formed by the joint Bayesian sparse model; they contain much less clutter and reveal the target with strong intensities. The TCRs of the composite scene images formed by the conventional CS and the proposed Bayesian approach are, respectively, 23 dB and 35.58 dB. This imaging improvement is because the proposed approach incorporates the joint-channel correlations into the imaging model, and thereby providing more prior knowledge for the imaging problem.

Figure 6.10 illustrates the ROC curves of the different imaging models. This

figure shows that by reconstructing multiple polarimetric images jointly and exploiting the inter-channel dependencies, target detection rate is significantly enhanced, compared with the methods that form the images individually for each channel.
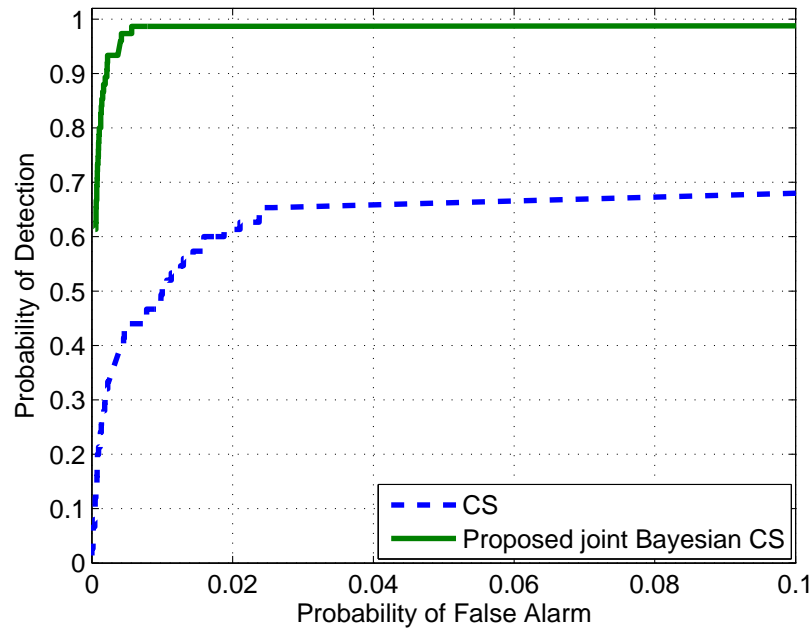
Figure 6.10: ROC curves of target detection on composite images formed by the standard CS model and the proposed joint Bayesian CS model.

6.5 Chapter summary

This chapter presented a novel approach for compressed multi-channel TWRI using joint Bayesian sparse representation. It combines multiple compressed measurement sets acquired from multi-view or multi-polarization sensing to enhance indoor target detection. Using these measurement sets, a joint Bayesian sparse approximation is proposed for estimating simultaneously the signal coefficients of different antennas. Furthermore, a subspace-projection technique is applied directly to the recovered coefficients to segregate wall clutter from the target returns. The remaining coefficients are combined to form a multitask imaging model. A composite image of the scene is reconstructed using joint Bayesian sparse learning, taking inter-channel dependences into account. Experimental

results show that the proposed approach enhances the scene reconstruction and target detection, compared with conventional imaging methods that form images independently for each channel.

Chapter 7

Conclusion

Chapter contents

Reliable imaging of indoor scenes using a small fraction of the entire data volume is essential for several civil applications and military operations. Forming images with reduced data volume, however, has traditionally hindered high-resolution imaging due to the restrictions on both bandwidth and aperture. Problems related to TWRI include prolonged data collection, extending objects, strong wall reflections, shadowing effects and target obstructions. To relax constraints on signal sampling schemes and logistic difficulties in data acquisition, and to address such problems for enhancing TWRI, this research investigates techniques based on compressed sensing and joint Bayesian CS. Several imaging models are proposed and their efficiency is validated through experimental results. The proposed research has produced efficient solutions for enhancing the capabilities of TWRI systems. A behind-the-wall scene can be imaged at a faster speed with far reduced data samples and thereby decreasing the time of data acquisition and cost of operation. In this chapter, Section 7.1 summarizes the major research activities and contributions of this project, and Section 7.2 suggests research directions for future work.

7.1 Research summary

The research activities have been presented in several chapters of the thesis and the main contributions are summarized as follows.

- Chapter 2 presents a comprehensive review of compressed sensing, a new area of signal processing, that has attracted considerable research interests recently. The aim of CS is to reconstruct robustly and precisely signals from far fewer measurements than is required by the Nyquist-Shannon sampling theorem. This chapter examines the principles of the CS theory including signal model, sparsity, compressibility, and incoherence. It also considers the conditions for stable and precise signal recovery and investigates several numerical algorithms for CS reconstruction. Furthermore, we introduce a Bayesian perspective for CS, namely Bayesian compressed sensing. The Bayesian CS framework extends the standard CS model and is useful for practical applications because the Bayesian CS produces a sparse solution as a full posterior distribution, rather than a point estimate as in conventional CS solvers.

- Chapter 3 introduces the principles of high-resolution TWR systems for imaging behind-the-wall objects. It describes TWR signal model and a conventional beamforming technique for image formation. The conventional beamforming method uses a full data volume collected by a huge array antenna and a large bandwidth signal to generate a high-resolution image of the scene. This chapter also reviews two major CS techniques for TWR image formation and signal reconstruction using only a reduced measurement set. It follows by discussions highlighting the challenges faced with existing CS techniques and identifying the research gaps that are addressed in this research project.

- Chapter 4 proposes a two-stage CS-based approach for TWR image for-

mation. The proposed approach is motivated by the fact that the sparsity assumption of the scene is usually not satisfied in practical applications due to multipath effects, wall returns, and large objects. The proposed algorithm includes two stages. In the first stage, an additional Gabor dictionary is incorporated to sparsely represent the scene. The use of the sparsifying dictionary increases the sparsity and thus enhances the CS recovery accuracy and reduces the number of measurements needed for scene reconstruction. In the second stage, the recovered data is used to form the scene image by employing the backprojection algorithm. Experimental results on both simulated and real data are presented which demonstrate the effectiveness of the proposed approach.

- Chapter 5 proposes a Bayesian scene reconstruction approach for compressed TWR sensing where not all the same frequencies are available along antennas. The Bayesian sparse model is used to jointly reconstruct TWR signal coefficients by exploiting both the intra-signal sparsity structure and the inter-signal correlations. An effective subspace-projection technique is applied directly to the recovered coefficients to suppress the wall clutter and reverberations. For efficient Bayesian scene reconstruction, a compact linear model is developed using PCA technique, which further compresses the measurement vector and dictionary. The experimental results show that the proposed approach enhances significantly the reconstruction accuracy, image quality in terms of target-to-clutter ratio and target detection rate. More importantly, the proposed imaging model can detect the targets even when the measurements are drastically reduced or several data samples are missing or corrupted.

- Chapter 6 addresses the problem of multi-channel TWRI using joint Bayesian compressed sensing. Multi-view or multi-polarization TWRI enhances the target detection and localization by combining multiple data sets acquired

from several sensing vantage points or channels. Collecting data at several vantage points improves imaging visibility, but also leads to prolonged data acquisition time, complex computation, and expensive hardware. To tackle such challenges and improve target detection, we propose a joint Bayesian CS model for multi-channel TWRI. The proposed approach reconstructs all the antenna signal coefficients simultaneously by exploiting the signal sparsity and inter-signal correlations, followed by subspace projection technique for wall clutter removal. A multitask linear imaging model is formulated including a composite scene image and images associated with different channels. All these scene images are jointly reconstructed by the joint Bayesian sparse approximation, taking the inter-channel dependencies into account. Experimental results using multi-view and multi-channel radar data demonstrate that the proposed approach improves reconstruction accuracy and produces a composite scene image in which the targets are enhanced and the background clutter is attenuated.

7.2 Future research directions

Using compressed sensing and Bayesian CS, this research project has developed efficient approaches for enhancing TWRI and urban sensing. The proposed approaches have been validated via extensive experiments. The results show they enhance the image quality and resolution, improve target detection and localization, and reduce the sampling rates and data acquisition times. Following the investigations presented in this dissertation, here is a list of possible research directions for future work.

- Enhancing the sparse representation with learning dictionary: In the proposed approaches, the scene image and antenna signals are sparsely represented using dictionaries with fixed atoms, such as Gabor, Fourier, wavelets, or DPSS. Dictionaries with their atoms learned from training data may be beneficial to scene representation and reconstruction. The training data can

be collected through long-term sensing surveillance. Dictionary learning techniques can potentially improve TWRI performances due to compact sparse representation and concise basis selection. Furthermore, learning dictionary approaches may enable us to model the inaccuracy and uncertainty caused by unknown wall parameters or radar positions in the problem of multi-view radar image reconstruction and multipath exploitation.

- Optimizing compressed TWR sensing: Compressed frequency measurements are acquired randomly for scene image formation and target detection. Because in TWR sensing, the RCS of targets depends on the frequency and the viewing angles, an optimized compressed sensing scheme, which automatically selects the measurements containing strong target reflections, can enhance the proposed approaches and reduce the number of data samples and the cost of processing.

- Separating wall returns and target signals directly with reduced data volume: In TWRI, the wall reflections reside in a low-rank subspace, and the target signals are sparse. Combining the prior knowledge of low-rank and sparsity may enhance the separation of the wall and target contributions even if several measurements are corrupted or missing. Furthermore, this approach may be useful for the reconstruction of scenes with moving targets since the motion can be captured by the sparse part of the received signals.

- Classifying targets directly in the compressed signal domain: The proposed approaches and other existing CS-based TWRI techniques aim to form high-quality scene images and enhance image-based target detection and localization. Developing a CS-based signal processing model that directly performs feature extraction and then classifies targets in the compressed signal domain may enhance the capabilities of real-time TWR sensing systems.

References

[1] F. Ahmad, M. G. Amin, and S. A. Kassam, "Synthetic aperture beamformer for imaging through a dielectric wall," *IEEE Trans. Aerospace and Electronic Systems*, vol. 41, no. 1, pp. 271–283, 2005.

[2] M. G. Amin (Ed.), *Through-The-Wall Radar Imaging*. Boca Raton, FL: CRC Press, 2010.

[3] M. G. Amin and K. Sarabandi, "Special issue on remote sensing of building interior," *IEEE Trans. Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1267 – 1268, 2009.

[4] F. Ahmad, M. G. Amin, and G. Mandapati, "Autofocusing of through-the-wall radar imagery under unknown wall characteristics," *IEEE Trans. Image Processing*, vol. 16, no. 7, pp. 1785–1795, 2007.

[5] C. Debes, M. G. Amin, and A. M. Zoubir, "Target detection in single- and multiple-view through-the-wall radar imaging," *IEEE Trans. Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1349–1361, 2009.

[6] C. H. Seng, A. Bouzerdoum, M. G. Amin, and S. L. Phung, "Probabilistic fuzzy image fusion approach for radar through wall sensing," *IEEE Trans. Image Processing*, vol. 22, no. 12, pp. 4938–4951, 2013.

[7] V. H. Tang, A. Bouzerdoum, and S. L. Phung, "Two-stage through-the-wall radar image formation using compressive sensing," *SPIE Journal of Electronic Imaging*, vol. 22, no. 2, pp. 021 006–1–021 006–10, 2013.

[8] V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Multi-view indoor scene reconstruction from compressed through-wall radar measurements using a joint Bayesian sparse representation," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Brisbane, Australia, April 2015, pp. 2419–2423.

[9] V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Enhanced wall clutter mitigation for through-the-wall radar imaging using joint Bayesian sparse signal recovery," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Florence, Italy, May 2014, pp. 7804–7808.

[10] V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Multi-view TWRI scene reconstruction using a joint Bayesian sparse approximation model," in *Proc. SPIE Defence, Security and Sensing: Compressive Sensing*, vol. 9484, Maryland, USA, April 2015, pp. 9484–1–9484–12.

[11] V. H. Tang, A. Bouzerdoum, S. L. Phung, and F. H. C. Tivive, "Enhanced through-the-wall radar imaging using Bayesian compressive sensing," in *Proc. SPIE Defence, Security and Sensing: Compressive Sensing*, vol. 8717, Maryland, USA, May 2013, pp. 87170I–1–87170I–12.

[12] A. Bouzerdoum, F. H. C. Tivive, and V. H. Tang, "Multi-polarization through-the-wall radar imaging using joint Bayesian compressed sensing," in *Proc. IEEE Int. Conf. Digital Signal Processing*, Hong Kong, China, August 2014, pp. 783–788.

[13] F. H. C. Tivive, A. Bouzerdoum, and V. H. Tang, "Multi-stage compressed sensing and wall clutter mitigation for through-the-wall radar image formation," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, Coruna, Spain, June 2014, pp. 489–492.

[14] D. L. Donoho, "Compressed sensing," *IEEE Trans. Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[15] E. J. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[16] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[17] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[18] M. Elad, "Optimized projections for compressed sensing," *IEEE Trans. Signal Processing*, vol. 55, no. 12, pp. 5695–5702, 2007.

[19] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries," *IEEE Trans. Information Theory*, vol. 54, no. 5, pp. 2210–2219, 2008.

[20] H. Nyquist, "Certain topics in telegraph transmission theory," *Transactions of the American Institute of Electrical Engineers*, vol. 47, no. 2, pp. 617–644, 1928.

[21] C. E. Shannon, "Communication in the presence of noise," *Proc. Institute of Radio Engineers*, vol. 37, no. 1, pp. 10–21, 1949.

[22] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 4, pp. 539–550, 1999.

[23] R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, 2007.

[24] Y. C. Eldar and G. Kutyniok (Ed.), *Compressed Sensing: Theory and Applications*. Cambridge University Press, Cambridge, UK, 2012.

[25] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Trans. Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[26] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, 2010.

[27] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*. Academic Press, Orlando, FL, 2008.

[28] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. New York: Springer, NY, 2010.

[29] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.

[30] D. G. Luenberger, *Optimization by Vector Space Methods*. John Wiley & Sons, New York, 1968.

[31] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

[32] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution," *Communications on Pure and Applied Mathematics*, vol. 59, no. 6, pp. 797–829, 2006.

[33] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[34] D. L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Information Theory*, vol. 47, no. 7, pp. 2845–2862, 2001.

[35] M. B. Wakin, "Compressive sensing fundamentals," in *Compressive Sensing for Urban Radar*. M. G. Amin (Ed.), Boca Raton, FL: CRC Press, Aug. 2014, pp. 2–30.

[36] B. K. Natarajan, "Sparse approximate solutions to linear systems," *SIAM Journal on Computing*, vol. 24, no. 2, pp. 227–234, 1995.

[37] A. M. Tillmann and M. E. Pfetsch, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," *IEEE Trans. Information Theory*, vol. 60, no. 2, pp. 1248–1259, 2014.

[38] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1999.

[39] E. J. Candes and J. Romberg, "ℓ_1-magic : Recovery of sparse signals via convex programming," *[Online]: http://users.ece.gatech.edu/ justin/l1magic/*, 2005.

[40] B. Efron, T. Hastie, L. Johnstone, and R. Tibshirani, "Least angle regression," *The Annals of Statistics*, vol. 32, no. 2, pp. 407–451, 2004.

[41] M. R. Osborne, B. Presnell, and B. A. Turlach, "On the LASSO and its dual," *Journal of Computational and Graphical Statistics*, vol. 9, no. 2, pp. 319–337, 2000.

[42] E. V. D. Berg and M. P. Friedlander, "Probing the Pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–23, 2008.

[43] J. Yang and Y. Zhang, "Alternating direction algorithms for ℓ_1-problems in compressive sensing," *SIAM Journal on Scientific Computing*, vol. 33, no. 1, pp. 250–278, 2011.

[44] Y. Tsaig and D. L. Donoho, "Extensions of compressed sensing," *Signal Processing*, vol. 86, no. 3, pp. 549–571, 2006.

[45] S. Becker, J. Bobin, and E. J. Candes, "NESTA: A fast and accurate first-order method for sparse recovery," *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 1–39, 2011.

[46] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *Journal of Machine Learning Research*, vol. 11, pp. 19–60, 2010.

[47] R. Ward, "Compressed sensing with cross validation," *IEEE Trans. Information Theory*, vol. 55, no. 12, pp. 5773–5782, 2009.

[48] P. Boufounos, M. F. Duarte, and R. G. Baraniuk, "Sparse signal reconstruction from noisy compressive measurements using cross validation," in *Proc. IEEE/SP Workshop on Statistical Signal Processing*, WI, USA, Aug. 2007, pp. 299–303.

[49] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[50] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Trans. Information Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.

[51] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[52] D. P. Wipf and B. D. Rao, "An empirical Bayesian strategy for solving the simultaneous sparse approximation problem," *IEEE Trans. Signal Processing*, vol. 55, no. 7, pp. 3704–3716, 2007.

[53] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Trans. Signal Processing*, vol. 52, no. 8, pp. 2153–2164, 2004.

[54] S. Ji, D. Dunson, and L. Carin, "Multitask compressive sensing," *IEEE Trans. Signal Processing*, vol. 57, no. 1, pp. 92–106, 2009.

[55] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Processing*, vol. 56, no. 6, pp. 2346–2356, 2008.

[56] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Bayesian compressive sensing using Laplace priors," *IEEE Trans. Image Processing*, vol. 19, no. 1, pp. 53–63, 2010.

[57] S. Rangan, A. K. Fletcher, and V. K. Goyal, "Asymptotic analysis of MAP estimation via the replica method and applications to compressed sensing," *IEEE Trans. Information Theory*, vol. 58, no. 3, pp. 1902–1923, 2012.

[58] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B*, vol. 58, no. 1, pp. 267–288, 1996.

[59] J. M. Bernardo and A. F. M. Smith, *Bayesian Theory*.   Wiley, New York, 1994.

[60] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis, 2nd ed*.   Boca Raton, FL: CRC Press, 2003.

[61] A. C. Faul and M. E. Tipping, "Analysis of sparse Bayesian learning," in *Proc. Advances in Neural Information Processing Systems*.   MIT Press, 2001, pp. 383–389.

[62] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.

[63] M. E. Tipping and A. C. Faul, "Fast marginal likelihood maximisation for sparse Bayesian models," in *Proc. Inter. Workshop on Artificial Intelligence and Statistics*, Key West, FL, Jan. 2003, pp. 3–6.

[64] M. A. Richards, *Fundamentals of Radar Signal Processing, 2nd ed*.   McGraw-Hill, New York, 2014.

[65] Y.-S. Yoon and M. G. Amin, "High-resolution through-the-wall radar imaging using beamspace MUSIC," *IEEE Trans. Antennas and Propagation*, vol. 56, no. 6, pp. 1763–1774, 2008.

[66] F. Ahmad, M. G. Amin, and S. A. Kassam, "Through-the-wall wideband synthetic aperture beamformer," in *Proc. IEEE Antennas and Propagation Society International Symposium*, vol. 3, CA, USA, June 2004, pp. 3059–3062.

[67] F. Ahmad, Y. Zhang, and M. G. Amin, "Three-dimensional wideband beamforming for imaging through a single wall," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 2, pp. 176–179, 2008.

[68] Y.-S. Yoon and M. G. Amin, "Imaging of behind the wall targets using wideband beamforming with compressive sensing," in *Proc. IEEE/SP Workshop on Statistical Signal Processing*, Cardiff, UK, Sep. 2009, pp. 93–96.

[69] F. Ahmad and M. G. Amin, "Noncoherent approach to through-the-wall radar localization," *IEEE Trans. Aerospace and Electronic Systems*, vol. 42, no. 4, pp. 1405–1419, 2006.

[70] M. G. Amin and F. Ahmad, "Wideband synthetic aperture beamforming for through-the-wall imaging," *IEEE Signal Processing Magazine*, vol. 25, no. 4, pp. 110–113, 2008.

[71] G. Wang and M. G. Amin, "Imaging through unknown walls using different standoff distances," *IEEE Trans. Signal Processing*, vol. 54, no. 10, pp. 4015–4025, 2006.

[72] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, S. Ting, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 83–91, 2008.

[73] P. Nagesh and L. Baoxin, "Compressive imaging of color images," in *Proc.*

IEEE Int. Conf. Acoustics, Speech and Signal Processing, Taipei, Taiwan, April 2009, pp. 1261–1264.

[74] J. Haupt and R. Nowak, "Compressive sampling Vs. conventional imaging," in *Proc. IEEE Int. Conf. Image Processing*, Atlanta, GA, Oct. 2006, pp. 1269–1272.

[75] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[76] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing MRI," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 72–82, 2008.

[77] J. Provost and F. Lesage, "The application of compressed sensing for photoacoustic tomography," *IEEE Trans. Medical Imaging*, vol. 28, no. 4, pp. 585–594, 2009.

[78] D. Malioutov, M. Cetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Processing*, vol. 53, no. 8, pp. 3010–3022, 2005.

[79] R. Mignot, L. Daudet, and F. Ollivier, "Room reverberation reconstruction: interpolation of the early part using compressed sensing," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 21, no. 11, pp. 2301–2312, 2013.

[80] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Processing*, vol. 57, no. 3, pp. 993–1009, 2009.

[81] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-nyquist sampling of sparse wideband analog signals," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 375–391, 2010.

[82] R. Baraniuk and P. Steeghs, "Compressive radar imaging," in *Proc. IEEE Radar Conf.*, Boston, MA, April 2007, pp. 128–133.

[83] X.-C. Xie, "Application of compressive sensing in 3D radar imaging," in *Proc. IEEE Int. Conf. Signal Processing*, Beijing, China, Oct. 2010, pp. 2117–2120.

[84] L. C. Potter, E. Ertin, J. T. Parker, and M. Cetin, "Sparsity and compressed sensing in radar imaging," *Proceedings of IEEE*, vol. 98, no. 6, pp. 1006–1020, 2010.

[85] M. T. Alonso, P. L. Pez-Dekker, and J. J. Mallorqui, "A novel strategy for radar imaging based on compressive sensing," *IEEE Trans. Geoscience and Remote Sensing*, vol. 48, no. 12, pp. 4285–4295, 2010.

[86] W. Shao, A. Bouzerdoum, and S. L. Phung, "Sparse representation of GPR traces with application to signal classification," *IEEE Trans. Geoscience and Remote Sensing*, vol. 51, no. 7, pp. 3922–3930, 2013.

[87] A. C. Gurbuz, J. H. McClellan, and W. R. Scott, "A compressive sensing data acquisition and imaging method for stepped frequency GPRs," *IEEE Trans. Signal Processing*, vol. 57, no. 7, pp. 2640–2650, 2009.

[88] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Signal Processing*, vol. 57, no. 6, pp. 2275–2284, 2009.

[89] Y.-S. Yoon and M. G. Amin, "Compressed sensing technique for high-resolution radar imaging," in *Proc. SPIE: Signal Processing, Sensor Fusion, and Target Recognition XVII*, Orlando, FL, March 2008, pp. 69 681A–1–69 681A–10.

[90] M. G. Amin, F. Ahmad, and W. Zhang, "Target RCS exploitations in compressive sensing for through wall imaging," in *Proc. Int. Waveform Diversity and Design Conf.*, Niagara Falls, ON, Aug. 2010, pp. 150–154.

[91] Y.-S. Yoon and M. G. Amin, "Through-the-wall radar imaging using compressive sensing along temporal frequency domain," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Dallas, TX, March 2010, pp. 2806–2809.

[92] M. Leigsnering, C. Debes, and A. M. Zoubir, "Compressive sensing in through-the-wall radar imaging," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Prague, Czech Republic, May 2011, pp. 4008–4011.

[93] J. Yang, A. Bouzerdoum, F. H. C. Tivive, and M. G. Amin, "Multiple-measurement vector model and its application to through-the-wall radar imaging," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Prague, Czech Republic, May 2011, pp. 2672–2675.

[94] M. G. Amin and F. Ahmad, "Compressive sensing for through-the-wall radar imaging," *Journal of Electronic Imaging*, vol. 22, no. 3, pp. 030 901–1–030 901–21, 2013, 10.1117/1.JEI.22.3.030901.

[95] Q. Huang, L. Qu, B. Wu, and G. Fang, "UWB through-wall imaging based on compressive sensing," *IEEE Trans. Geoscience and Remote Sensing*, vol. 48, no. 3, pp. 1408–1415, 2010.

[96] Y.-S. Yoon and M. G. Amin, "Spatial filtering for wall-clutter mitigation in through-the-wall radar imaging," *IEEE Trans. Geoscience and Remote Sensing*, vol. 47, no. 9, pp. 3192–3208, 2009.

[97] M. Duman and A. C. Gurbuz, "Analysis of compressive sensing based through the wall imaging," in *Proc. IEEE Radar Conference*, Atlanta, GA, May 2012, pp. 0641–0646.

[98] C. H. Seng, A. Bouzerdoum, F. H. C. Tivive, and M. G. Amin, "Fuzzy logic-based image fusion for multi-view through-the-wall radar," in *Proc. Int.*

Conf. Digital Image Computing: Techniques and Applications, Sydney, NSW, Dec. 2010, pp. 423–428.

[99] F. Ahmad, "Multi-location wideband through-the-wall beamforming," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Las Vegas, NV, April 2008, pp. 5193–5196.

[100] F. Ahmad and M. G. Amin, "Multi-location wideband synthetic aperture imaging for urban sensing applications," *Journal of the Franklin Institute*, vol. 345, no. 6, pp. 618–639, 2008.

[101] K. M. Yemelyanov, N. Engheta, A. Hoorfar, and J. A. McVay, "Adaptive polarization contrast techniques for through-wall microwave imaging applications," *IEEE Trans. Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1362–1374, 2009.

[102] A. A. Mostafa, C. Debes, and A. M. Zoubir, "Segmentation by classification for through-the-wall radar imaging using polarization signatures," *IEEE Trans. Geoscience and Remote Sensing*, vol. 50, no. 9, pp. 3425–3439, 2012.

[103] S. Fischer, G. Cristobal, and R. Redondo, "Sparse overcomplete Gabor wavelet representation based on local competitions," *IEEE Trans. Image Processing*, vol. 15, no. 2, pp. 265–272, 2006.

[104] R. Fazel-Rezai and W. Kinsner, "Image decomposition and reconstruction using two-dimensional complex-valued Gabor wavelets," in *Proc. IEEE Int. Conf. Cognitive Informatics*, Lake Tahoo, CA, Aug. 2007, pp. 72–78.

[105] K. N. Chaudhury and M. Unser, "Construction of Hilbert transform pairs of wavelet bases and Gabor-like transforms," *IEEE Trans. Signal Processing*, vol. 57, no. 9, pp. 3411–3425, 2009.

[106] F. H. C. Tivive, A. Bouzerdoum, and M. G. Amin, "An SVD-based approach

for mitigating wall reflections in through-the-wall radar imaging," in *Proc. IEEE Radar Conf.*, Kansas City, MO, May 2011, pp. 519–524.

[107] E. Lagunas, M. G. Amin, F. Ahmad, and M. Najar, "Wall mitigation techniques for indoor sensing within the compressive sensing framework," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop*, Hoboken, NJ, June 2012, pp. 213–216.

[108] E. Lagunas, M. G. Amin, F. Ahmad, and M. Najar, "Joint wall mitigation and compressive sensing for indoor image reconstruction," *IEEE Trans. Geoscience and Remote Sensing*, vol. 51, no. 2, pp. 891 – 906, 2013.

[109] O. Scherzer, *Handbook of Mathematical Methods in Imaging.* New York: Springer Science, 2011.

[110] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Trans. Information Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[111] C. Debes, J. Hahn, A. M. Zoubir, and M. G. Amin, "Target discrimination and classification in through-the-wall radar imaging," *IEEE Trans. Signal Processing*, vol. 59, no. 10, pp. 4664–4676, 2011.

[112] M. Elad, P. Milanfar, and R. Rubinstein, "Analysis versus synthesis in signal priors," *Inverse Problems*, vol. 23, no. 3, pp. 947–968, 2007.

[113] C. H. Seng, A. Bouzerdoum, M. G. Amin, and S. L. Phung, "Two-stage fuzzy fusion with applications to through-the-wall radar imaging," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 4, pp. 687–691, 2012.

[114] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, "Signal processing with compressive measurements," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 445–460, Apr. 2010.

[115] V. M. Patel, G. R. Easley, D. M. Healy, Jr., and R. Chellappa, "Compressed synthetic aperture radar," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 244–254, Apr. 2010.

[116] A. Bouzerdoum, J. Yang, and F. H. C. Tivive, "Compressive sensing for multi-polarization through-the-wall radar imaging," in *Compressive Sensing for Urban Radar*. M. G. Amin (Ed.), Boca Raton, FL: CRC Press, Aug. 2014, pp. 232–248.

[117] F. Ahmad and M. G. Amin, "Through-the-wall human motion indication using sparsity-driven change detection," *IEEE Trans. Geoscience and Remote Sensing*, vol. 51, no. 2, pp. 881–890, 2013.

[118] F. Ahmad and M. G. Amin, "Partially sparse reconstruction of behind-the-wall scenes," in *Proc. SPIE Defence, Security and Sensing: Compressive Sensing*, vol. 8365, Maryland, USA, April 2012, pp. 83 650W–1–83 650W–9.

[119] F. H. C. Tivive and A. Bouzerdoum, "An improved SVD-based wall clutter mitigation method for through-the-wall radar imaging," in *Proc. IEEE Workshop on Signal Processing Advances in Wireless Communications*, Darmstadt, Germany, June 2013, pp. 430–434.

[120] F. H. C. Tivive, A. Bouzerdoum, and M. G. Amin, "A subspace projection approach for wall clutter mitigation in through-the-wall radar imaging," *IEEE Trans. Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2108–2122, 2015.

[121] D. Baron, M. B. Wakin, M. F. Duarte, S. Sarvotham, and R. G. Baraniuk, "Distributed compressed sensing," Rice University, TREE-0612, Tech. Rep., Houston, TX, Nov. 2006.

[122] F. Ahmad, J. Qian, and M. G. Amin, "Wall clutter mitigation using discrete prolate spheroidal sequences for sparse reconstruction of indoor stationary

scenes," *IEEE Trans. Geoscience and Remote Sensing*, vol. 53, no. 3, pp. 1549–1557, 2015.

[123] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. part I: Greedy pursuit," *Signal Processing*, vol. 86, no. 3, pp. 572–588, 2006.

[124] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Processing*, vol. 53, no. 7, pp. 2477–2488, 2005.

[125] J. A. Tropp, "Algorithms for simultaneous sparse approximation. part II: Convex relaxation," *Signal Processing*, vol. 86, no. 3, pp. 589–602, 2006.

[126] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[127] D. J. C. MacKay, "Bayesian interpolation," *Neural Computation*, vol. 4, no. 3, pp. 415–447, 1992.

[128] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 33, no. 2, pp. 387–392, 1985.

[129] M. A. Davenport and M. B. Wakin, "Compressive sensing of analog signals using discrete prolate spheroidal sequences," *Applied and Computational Harmonic Analysis*, vol. 33, no. 3, pp. 438–472, 2012.

[130] M. G. Amin, F. Ahmad, and W. Zhang, "A compressive sensing approach to moving target indication for urban sensing," in *Proc. IEEE Radar Conf.*, Kansas City, MO, May 2011, pp. 509–512.

[131] T. Dogaru and C. Le, "Through-the-wall small weapon detection based on polarimetric radar techniques," Army Research Laboratory, ARL-TR-5041, Tech. Rep., Adelphi MD, Dec. 2009.

[132] H. Griffiths and C. Baker, *Radar Imaging for Combatting Terrorism*, ser. NATO Security through Science Series, J. Byrnes, Ed. Springer Netherlands, 2007.

[133] Y. Zhang and M. G. Amin, "Joint Doppler and polarization characterization of moving targets," in *Proc. IEEE Antennas and Propagation Society International Symposium*, vol. 3, CA, USA, June 2004, pp. 3083–3086.

[134] Q. Wu, Y. D. Zhang, F. Ahmad, and M. G. Amin, "Compressive-sensing-based high-resolution polarimetric through-the-wall radar imaging exploiting target characteristics," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 1043–1047, 2015.

[135] W. Zhang, C. Thajudeen, and A. Hoorfar, "Polarimetric through-the-wall imaging," in *Proc. International Symposium on Electromagnetic Theory*, Berlin, Germany, Aug. 2010, pp. 471–474.

[136] C. Debes, A. M. Zoubir, and M. G. Amin, "Enhanced detection using target polarization signatures in through-the-wall radar imaging," *IEEE Trans. Geoscience and Remote Sensing*, vol. 50, no. 5, pp. 1968–1979, 2012.

[137] M. Thiel and K. Sarabandi, "Ultrawideband multi-static scattering analysis of human movement within buildings for the purpose of stand-off detection and localization," *IEEE Trans. Antennas and Propagation*, vol. 59, no. 4, pp. 1261–1268, 2011.

[138] J. Yang, A. Bouzerdoum, and M. G. Amin, "Multi-view through-the-wall radar imaging using compressed sensing," in *Proc. European Signal Processing Conference*, Aalborg, Denmark, Aug. 2010, pp. 1429–1433.

[139] K. S. Kunz and R. J. Luebbers, *The Finite Difference Time Domain Method for Electromagnetics*. Boca Raton, FL: CRC Press, 1993.