A Survey on the Low-Dimensional-Model-based Electromagnetic Imaging

Lianlin Li¹, Martin Hurtado², Feng Xu³, Bing Chen Zhang⁴, Tian Jin⁵, Tie Jun Cui⁶, Marija Nikolic Stevanovic⁷ and Arye Nehorai⁸

¹ Peking University; lianlin.li@pku.edu.cn
² National University of La Plata; martin.hurtado@ing.unlp.edu.ar
³ Fudan University; fengxu@fudan.edu.cn
⁴ Institute of Electronics, Chinese Academy of Sciences; bczhang@mail.ie.ac.cn
⁵ National University of Defense Technology; tianjin@nudt.edu.cn
⁶ Southeast University; tjcui@seu.edu.cn
⁷ University of Belgrade; mnikolic@etf.rs
⁸ Washington University in St. Louis; nehorai@ese.wustl.edu

ABSTRACT

The low-dimensional-model-based electromagnetic imaging is an emerging member of the big family of computational imaging, by which the low-dimensional models of underlying signals are incorporated into both data acquisition systems and reconstruction algorithms for electromagnetic imaging, in order to improve the imaging performance and break the bottleneck of existing electromagnetic imaging methodologies. Over the past decade, we have witnessed profound impacts of the low-dimensional models on electromagnetic imaging. However, the low-dimensional-model-based electromagnetic imaging remains at its early stage, and many

important issues relevant to practical applications need to be carefully investigated. Especially, we are in the big-data era of booming electromagnetic sensing, by which massive data are being collected for retrieving very detailed information of probed objects. This survey gives a comprehensive overview on the low-dimensional models of structure signals, along with its relevant theories and low-complexity algorithms of signal recovery. Afterwards, we review the recent advancements of low-dimensional-model-based electromagnetic imaging in various applied areas. We hope this survey could bridge the gap between the model-based signal processing and the electromagnetic imaging, advance the development of low-dimensional-model-based electromagnetic imaging, and serve as a basic reference in the future research of the electromagnetic imaging across various frequency ranges.
Electromagnetic imaging has been a powerful technique in various civil and military applications across medical imaging, geophysics, space exploration, resources and energy survey, etc., where the operational frequency ranges from the very low frequency (like tens of Hertz) through microwave, millimeter wave, and Terahertz, up to optical frequencies Devaney, 2012; Pastorino, 2010; Born and Wolf, 1980; Zhdanov, 2009; Kunitsyn and Tereshchenko, 2003; Solimene et al., 2014; Fornaro et al., 2014. The electromagnetic imaging problem or the electromagnetic inverse scattering problem consists of determining the unknown features (including geometrical and physical parameters) of an object from processing measured electromagnetic data Pastorino, 2010; Zhdanov, 2009. Essentially, it is strongly nonlinear and ill-posed due to the complicated interaction between the electromagnetic wavefield and the imaging scene Pastorino, 2010; Zhdanov, 2009. In principle, this problem could be addressed by employing nonlinear iterative optimization methods, but these iterative methods are computationally prohibitive even for the moderate-scale problem Pastorino, 2010. In practice, one resorts to the linearized approximate solution to the rigorous inverse scattering problem, for example, Born-approximation Pastorino, 2010;
Born and Wolf, 1980. Nonetheless, the resulting inverse problem remains notoriously ill-posed since the measurements available are inadequate typically compared to the unknowns to be retrieved. Especially, there are increasing continuously demands on the imaging resolution of detailed information of probed object nowadays, which broaden the gap between the unknowns of interest and the measurements available further. Moreover, the measurements are noisy and suffer from unknown ambiguous parameters, which makes the electromagnetic imaging problem more challenging.

Put formally, the electromagnetic imaging can be formulated as $y = A_\theta(x) + n$, where the quantity $y$ indicates the vectorized measurements corrupted with additive noise $n$, $x$ denotes the unknown (e.g., the reflectivity of imaged scene) to be retrieved, $A_\theta$ is a mapping operator with the subscript $\theta$ highlighting possible unknown ambiguous parameters Devaney, 2012; Pastorino, 2010; Born and Wolf, 1980; Zhidanov, 2009; Kunitsyn and Tereshchenko, 2003; Solimene et al., 2014.

As argued above, the nonlinear inverse scattering, i.e. $A_\theta$ being nonlinear, is limited to the small-scale problem due to its very expensive computational cost. For this reason, we are restricted ourselves into the case of $A_\theta$ being linear. Furthermore, without the loss of generality, we assume no ambiguous parameters involved, implying that the subscript $\theta$ vanishes. Consequently, $A_\theta$ becomes $A$. Then, the electromagnetic imaging problem consists of retrieving the unknowns $x$ from the noisy measurements $y$. In probabilistic framework, the estimation of $x$ amounts to evaluating the posterior probability of $x$ conditional on $y$ Tarantola, 2005; Idier, 2013; Tveito et al., 2010:

$$Pr(x|y) = \frac{1}{Z} Pr(y|x) Pr(x)$$  \hspace{1cm} (1.1)

where $Z = \int Pr(y|x) Pr(x) dx$ is the normalized factor (or the partition function), $Pr(y|x)$ is the likelihood function of $x$, and $Pr(x)$ is the prior knowledge on $x$. Once obtaining the posterior probability $Pr(x|y)$, we can numerically or analytically calculate desirable statistical quantities. We are particularly interested in the maximum a posteriori (MAP)
mode denoted by \(x_{\text{MAP}}\) \cite{Tarantola2005,Idier2013,Tveito2010}:

\[
x_{\text{MAP}} = \arg \max_x \log \Pr (x|y)
= \arg \max_x \left[ \log \Pr (y|x) + \log \Pr (x) \right] \quad (1.2)
= \arg \min_x \left[ \frac{1}{2} \|y - A(x)\|_2^2 - \log \Pr (x) \right] \quad (1.3)
\]

As pointed out previously, Eq. 1.3 is ill-posed due to the inadequate measurements compared to the unknowns to be retrieved, especially when the finer details of probed scene are desirable. Since the measurements are incomplete, an infinite number of solutions, however, being non-meaningful, match measurements. Therefore, one crucial task of electromagnetic imaging is to select the most meaningful solution out of the potential solutions. In terms of Bayesian analysis, the model function \(\Pr (x)\) provides the speculative knowledge on the underlying signal \(x\), which is, if correct, helpful in suppressing remarkably the solution uncertainty by complementing the incomplete measurements. In this sense, one feasible approach to the above task is the exploration of the correct model \(\Pr (x)\) in the design of imaging systems and imaging algorithms.

Most of real-world signals have low-dimensional models, known as being of the structured sparsity or structured compressibility \cite{Baraniuk2010,Elad2010,Davenport2012,Bach2011}. Here, we mean by structure that a transformed domain or manifold, being either deterministic or probabilistic, exists such that over which the transformed coefficients are sparse or compressible. By sparsity, we mean that a signal of length \(n\) has \(k \ll n\) nonzero elements; in contrast, we mean by compressibility that a signal of length \(n\) can be approximated with certain accuracy by a signal with only \(k \ll n\) nonzero coefficients \cite{Baraniuk2010b}. Low-dimensional signal models affect significantly the data acquisition, analysis and later processing \cite{Elad2012,Wainwright2014,Duarte2011}, which has profoundly broken bottlenecks set by the well-known Nyquist-Shannon theory founded by Kotelnikov, Nyquist, Shannon, Whittaker \emph{et al}. A celebrated theory known as compressive sensing, founded by Candès, Tao, Romberg, Donoho \emph{et al}., states that the sparse or compressible signal can be
Introduction

accurately and efficiently retrieved from its low-dimensional projections Candès and Tao, 2006; Candès and Tao, 2005; Candès et al., 2004; Candès et al., 2006; Donoho, 2006a. Afterwards, many more realistic and richer low-dimensional signal models along with the guarantee of theories and algorithms have been discovered and investigated, which affect the data acquisition, analysis and processing significantly Duarte and Eldar, 2011; Baraniuk et al., 2010a; He and Carin, 2009; He et al., 2010; Eldar et al., 2010. For instance, for tree- and block-structured signals, Baraniuk et al. established the theory of model-based compressive sensing (CS) along with reconstruction algorithms He and Carin, 2009.

Low-dimensional-model based signal processing (model-based SP, for short) differs from the Nyquist-Shannon theory based methods (conventional SP, for short) in several important aspects, as summarized in Table 1.1. As opposed to the conventional SP, which only uses the information provided by the measurement, model-based SP uses both the measurement and the prior knowledge, enabling us to break the limits in classical signal processing. For example, the number of measurements required by the Nyquist-Shannon theory can be dramatically reduced He and Carin, 2009; He et al., 2010; Eldar et al., 2010; Bevacqua et al., 2017, the Rayleigh resolution limit can be readily beat Candès and Fernandez-Granda, 2013; Candès and Fernandez-Granda, 2012, and so on. In addition, these two frameworks differ in the manner in which they deal with signal recovery. For conventional SP, the signal recovery is accomplished through simple sinc interpolation with marginal computational cost Davenport et al., 2012. In contrast, the model-based signal recovery is achieved by implementing nonlinear iterative algorithms, which needs apparently expensive computational resources He and Carin, 2009; He et al., 2010; Eldar et al., 2010.

Over past years, we have witnessed the notable impacts of low-dimensional-model-based signal processing, more strictly, the sparse-model-based signal processing, on the electromagnetic imaging Candès et al., 2004; Massa and Teixeira, 2015; Massa et al., 2015; Potter et al., 2010; Hunt and Smith, 2013; Li et al., 2015; Stevanovic et al., 2016a; Anselmi et al., 2015; Scapaticci et al., 2014; Bevacqua et al., 2015; Winters et al., 2010; Hurtado et al., 2011; Solimene et al., 2012;
Table 1.1: Comparisons between two frameworks of conventional and low-dimensional-model-based signal processing.

<table>
<thead>
<tr>
<th></th>
<th>Conventional SP</th>
<th>Model-based SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data input</td>
<td>Measurements</td>
<td>Measurements plus prior knowledge</td>
</tr>
<tr>
<td>Math. Model $y = Ax$</td>
<td>$\dim (y) \geq \dim (x)$</td>
<td>$\dim (y) \ll \dim (x)$</td>
</tr>
<tr>
<td>Data acquisition</td>
<td>Nyquist sampling rate</td>
<td>RIP criterion</td>
</tr>
<tr>
<td>Data communication</td>
<td>Sampling followed by compression</td>
<td>Compressed sampling</td>
</tr>
<tr>
<td>Representation</td>
<td>Independent pixels</td>
<td>Structure or pattern</td>
</tr>
<tr>
<td>Statistical model</td>
<td>Gaussian</td>
<td>Non-Gaussian</td>
</tr>
<tr>
<td>Signal Recovery</td>
<td>Sinc interpolation</td>
<td>Nonlinear optimization</td>
</tr>
<tr>
<td>Resolution</td>
<td>Rayleigh limit</td>
<td>Better than Rayleigh limit</td>
</tr>
</tbody>
</table>

Soldovieri et al., 2012; Oliveri et al., 2011; Baraniuk and Steeghs, 2007; Cetin and Karl, 2001; Cetin et al., 2004; Huang et al., 2010; Bevacqua et al., 2017. Imaging techniques, that exploit low-dimensional-model of the underlying scene, are becoming more and more popular thanks to their ability to mitigate the theoretical and practical difficulties arising in the associated inverse problem, while properly complying with several common applicative requirements (e.g., reduced computational costs, high spatial resolution, and robustness to the noise). Such an increased interest is proved by a vast of publications in several areas (e.g., electromagnetic inverse scattering, radar, microwave imaging, and array synthesis), and special sessions in relevant international conferences, as well as special issues in leading-edge journals. Invoked by the concept of compressive sensing, Hunt et al. Hunt and Smith, 2013 and Li et al. Li et al., 2015 invented the single-pixel images for microwave imaging of sparse scenes, demonstrating the important potential of sparse-signal model of imaged scene in developing the apparatus of low-complexity and low-cost data acquisition. By now, the most affected issue related to electromagnetic imaging is the development
of low-order model-based imaging algorithms. For instance, there are a large amount of efforts have been made to melt the sparsity-promoted regularization with the iterative algorithms of electromagnetic inverse scattering, being different in the selection of sparsifying transformations, leading to the feature-enhanced reconstruction. As argued above, the non-linear electromagnetic inverse scattering is limited to the small-scale problem due to very expensive computational complexity. For this reason, we leave the discussion about them out of this survey, and refer to Stevanovic et al., 2016a; Anselmi et al., 2015; Scapaticci et al., 2014; Bevacqua et al., 2015; Winters et al., 2010 for detailed discussions. For the linearized electromagnetic imaging (Born-based tomography Hurtado et al., 2011; Solimene et al., 2012; Soldovieri et al., 2012; Oliveri et al., 2011 and signal-based radar imaging Baraniuk and Steeghs, 2007; Cetin and Karl, 2001; Cetin et al., 2004; Huang et al., 2010; Yoon and Amin, 2010; Leigsnering et al., 2013; Zhu and Bamler, 2014; Zhang et al., 2011a), the sparse-model (or compressible-model) of imaged scene have been exploited, and a large amount of low-order model-based imaging algorithms have been developed, demonstrating that the usefulness of sparse-model in enhancing the image quality and reducing the number of measurements. It is really appealing to incorporate low-dimensional models of underlying signals into the electromagnetic imaging, in order to reduce the number of measurements, improve the imaging resolution, enhance the capability of object recognition and classification, and so on. For this reason, we refer to this methodology of electromagnetic imaging as the low-dimensional-model-based electromagnetic imaging (model-based electromagnetic imaging, for short). Here, we would like to give a formal description about it:

**Definition 1:** The low-dimensional-model-based electromagnetic imaging is the object-oriented and feature-enhanced electromagnetic imaging methodology by incorporating the knowledge of the structured models of underlying signals into the data acquisition system and the reconstruction algorithm, in order to reducing the number of measurements, improve the imaging resolution, enhance the capability of object recognition and classification, and so on.
Although the model-based signal processing by itself has arrived at relatively mature level with a solid body of theories Duarte and Eldar, 2011; Candès and Tao, 2006; Candès and Tao, 2005; Candès et al., 2004; Candès et al., 2006; Donoho, 2006a; Baraniuk et al., 2010a; He and Carin, 2009; He et al., 2010; Eldar et al., 2010; Candès and Fernandez-granda, 2013; Candès and Fernandez-Granda, 2012; Gribonval et al., 2012; Candès and Recht, 2012; Candès and Plan, 2009; Donoho and Tanner, 2010; Donoho, 2006b; Donoho and Elad, 2003; Donoho and Elad, 2006; Donoho et al., 2006; Donoho and Tanner, 2005 and algorithms Donoho and Tsaig, 2008; Tibshirani, 1996; Chen et al., 1998; Koh et al., 2007; Zou and Hastie, 2005; Candès et al., 2007; Mallat and Zhang, 1993; Needell and Tropp, 2009; Nesterov, 2013; Yuan and Lin, 2006; Tibshirani et al., 2005; Van Den Berg et al., 2008; Friedman et al., 2010; Tipping, 2001; Ji et al., 2008, its interactions with electromagnetic imaging remains challenging and many important issues are deserved to be studied in-depth. The model-based electromagnetic imaging focuses on four major aspects as following.

First, it is desirable the development of the next-generation imaging system with low-cost, low-complexity and high efficiency by the way of the optimal design of the waveform, the programmable or reconfigurable antenna, the configuration of sparse sensor array, etc. For instance, the establishment of novel compressive radar is appealing for ultra-wideband (UWB) radar imaging Huang et al., 2010; Yoon and Amin, 2010, since it is really too costly, or even physically impossible, to build devices capable of acquiring samples at the necessary rate in the context of classical signal processing.

Second, it is desirable to establish an easy-implementation of imaging formulations, which account for the real interaction between the electromagnetic wavefield and the probed scene. The interaction between the electromagnetic wavefield and the probed scene is nonlinear in essence; however, most of mathematical formulations of electromagnetic imaging considered so far are linear, failing to capture fully the undergoing physical mechanism in some practical cases Tsang et al., 1985; Bucci and Isernia, 2016; Colton and Kress, 2012. Nikolic et al. attempted to link the electromagnetic mechanism and the sparse-signal model by using the
equivalent electromagnetic currents, and demonstrated that with this model, the sparseness of imaged scene supports the reconstruction of non-convex shape of 2D PEC targets Nikolic et al., 2013. However, such methodology is limited to the single-frequency imaging configuration, since the equivalent currents vary with the operational frequencies.

Third, it is desirable to discover more realistic and richer low-dimensional models of the underlying electromagnetic information. By now, the low-dimensional models utilized in the area of electromagnetic imaging are nearly simple sparse or compressible models; more realistic and richer models are not fully investigated.

Fourth, we are in the deluge of massive electromagnetic data coming from the continuously increasing demands on retrieving very detailed information of objects nowadays. Therefore, it is crucial to develop efficient reconstruction algorithms for treating massive measurements and high-dimensional variables. From the aspect of computational complexity, it is also important to develop algorithmic frameworks trading off the imaging accuracy with the computational cost.
2

Progress on Model-based Signal Processing

2.1 Low-dimensional Signal Models

A low-dimensional signal model is a powerful tool in several aspects including saving the number of measurements, reducing the effective number of permitted subspaces, deeper dimensionality reduction, meaningful interpretation, and so on. A signal denoted by an \( N \)-length vector \( x \) is said to have low-dimensional model if there is a transform \( D : \mathbb{R}^L \rightarrow \mathbb{R}^N \) such that the \( L \)-length vector \( \alpha \), being the solution of \( \{ \alpha, \| x - D \alpha \|_2^2 \leq \varepsilon \} \), is the \((K, q)\)-compressible vector, where \( K \ll N \). Moreover, a \( L \)-length vector \( \alpha \) is said to be \((K, q)\)-compressible if \( \bar{\sigma}_K(\alpha)_q = \inf_{\| \beta \|_0 \leq K} \| \beta - \alpha \|_q \ll 1 \) Gribonval et al., 2012 (with slight modification).

Herein, \( \bar{\sigma}_K(\alpha)_q = \inf_{\| \beta \|_0 \leq K} \| \beta - \alpha \|_q \) is the \( K \)-term best approximation error of \( \alpha \) in the \( q \)-norm sense, \( \| \alpha \|_q (0 < q < \infty) \) is the \( q \)-(quasi)norm of \( \alpha \), and \( \| \beta \|_0 \) counts the non-zero elements of \( \beta \). Parallel to this deterministic model of compressible signals, Gribonval et al. introduced the concept of compressible distribution as following. A probabilistic distribution \( p(x) \) is said to be \((K, q)\)-compressible if its i.i.d. samples are almost \((K, q)\)-compressible Gribonval et al., 2012. Besides the Bernoulli-Gaussian distribution as sparse distribution, other popular compressible distributions includes the t-student, the generalized Pareto,
the log-logistic, and so on Baraniuk et al., 2010b; Gribonval et al., 2012; Tipping, 2001; Lavielle, 1993; Vila and Schniter, 2011. Here, it worthwhile pointing out that although some probabilistic distributions are not compressible, its use in combination with the MAP estimation indeed promotes the compressible solution Baraniuk et al., 2010b; Gribonval et al., 2012; a typical example is the Laplacian distribution.

The simplest case is that the elements of a compressible signal $\alpha$ are independent. In this model, the inter-structure among elements is ignored. It is well accepted, especially from the results in the community of machine learning, that there is more structures or patterns of undergoing signals. These richer structures, which cannot be captured by simple sparse models and regularizations, if exploited, can improve significantly the estimation of parameters as well as the interpretability of the estimates Baraniuk et al., 2010b; Bach et al., 2011; Elad, 2012; Duarte and Eldar, 2011; Baraniuk et al., 2010a; He and Carin, 2009; He et al., 2010; Eldar et al., 2010. Roughly speaking, there are three kinds of richer low-dimensional models:

2.1.1 Non-overlapping independent groups

This model suggests that the structure pattern manifests itself in the manner of sparse group-type (or block or clustered) structure instead of independent sparse elements Tibshirani et al., 2005; Van Den Berg et al., 2008; Friedman et al., 2010; Huang and Zhang, 2010; Zhang and Rao, 2012. One type of widespread structure consists of the non-overlapping independent groups; its formal definition can be described as:

**Definition 2:** A compressible vector $\alpha$ is assumed to be with non-overlapping independent groups if there exists a decomposition of $\alpha = \sum_i \alpha_i$ along with $\text{supp}(\alpha_i) \cap \text{supp}(\alpha_j) = \emptyset$ for $i \neq j$, where the elements of component $\{\alpha_i\}$ are bundled together.

In a common place, this structured sparsity is measured by the mixed-norm $l_{p/q}$, i.e., $\Omega(\alpha) = \left(\sum_i \|\alpha_i\|_q^p\right)^{1/p}$, where $q \geq 2$ and $0 \leq p \leq 1$ Tibshirani et al., 2005. The mixed-norm behaviors like
the $l_p$-norm (sparse inducing norm) of the re-organized new row vector $\tilde{\alpha} = [\|\alpha_1\|_q, \|\alpha_2\|_q, \ldots, \|\alpha_n\|_q]$. In this sense, the signal models, theories and algorithms for independent group structures are the immediate extension of the corresponding results for the structure with independent sparse elements. Solving the inverse problem regularized with independent group structured sparsity can be pursued with the deterministic algorithms, such as, group LASSO algorithms Friedman et al., 2010, group greedy algorithm Van Den Berg et al., 2008, etc. In parallel, some probabilistic models for independent group structures and associated reconstruction algorithms have been developed as well Peleg et al., 2012; Zhang and Rao, 2012; Amit et al., 2007. In machine learning community, the mixture of Gaussian Yu and Sapiro, 2011; Duarte-Carvajalino et al., 2012 is a very popular probabilistic model for characterizing group structure data.

The group structure poses dependencies of representation coefficients. Besides it, there are other ways to impose such structure on the representation vector and its support. Among them, the well-known TV (Total Variation) regularization is a promising tool for measuring piecewise smooth (or group or block) signals, where the dependence of neighbored elements are imposed by gradient operator. Furthermore, the non-local operator of generalized gradient operator developed by Gilboa and Osher plays the role similar to the TV measure Gilboa and Osher, 2008.

### 2.1.2 Overlapping groups

In the past years, there is a number of more realistic and richer low-dimensional models, generalizing above non-overlapping group structure to the overlapping groups Wainwright, 2014; Jacob et al., 2009; Jenatton et al., 2010; Bach et al., 2012; Shervashidze and Bach, 2015 and tree-structured groups Baraniuk et al., 2010a; He and Carin, 2009; He et al., 2010. See Bach et al., 2011; Bach et al., 2012 for a more detailed overview on this topic. Here, we adopt the description of overlapping group structure defined in Bach et al., 2012, i.e.
Definition 3: A compressible vector $\alpha$ is assumed to be with non-overlapping independent groups if there exists a decomposition of $\alpha = \sum_i \alpha_i$ along with $\text{supp}(\alpha_i) \cap \text{supp}(\alpha_j) \neq \emptyset$ for $i \neq j$.

This type of signals with low-dimensional structure could be measured by the mixed-norm $l_{p/q}$, i.e., $\Omega(\alpha) = \left( \sum_i w_g \| \alpha_i \|_q^p \right)^{1/p}$, where $q \geq 2$ and $0 \leq p \leq 1$ Bach et al., 2012. Generally, we do not have this prior knowledge about the relevance of individual groups, either for overlapping or non-overlapping groups. The problem of automatically choosing appropriate weights for groups of variables is an important open research problem in structured sparsity. Shervashidze and Bach proposed a framework for learning group relevance from data using probabilistic modeling with a broad family of heavy-tailed priors and derived a variational inference scheme to learn the parameters of these priors Shervashidze and Bach, 2015.

2.1.3 Learning dictionary

It should be emphasized that low-dimensional models above are defined over a known transform, such as discrete cosine transform (DCT), wavelet, curvelet, or other redundant dictionaries, etc. The analytical transformations like DCT and wavelet are too general to describe well the task-dedicated features of undergoing signals, especially for the application of object recognition where the subtle features need to be distinguished. Partially for this consideration, dictionary learning is a prominent part of structural signal processing, as it enables the extraction of an underlying structured dictionary from a given set of signal examples. Typical dictionary learning algorithms range from principal component analysis (PCA) Jolliffe, 1986, generalized PCA Vidal et al., 2005, and the K-singular value decomposition (K-SVD) Aharon et al., 2006. A comprehensive overview on this topic was made in Tosic and Frossard, 2011. Dictionary learning methods avoid choosing a fixed dictionary in which some atoms might be of limited use, and therefore offer refined dictionaries that adapt the structure of the undergoing task. Here, we would like to give a general definition of dictionary learning:
Definition 4: Give an $n$-samples training set $Y = [y_1, y_2, \ldots, y_n]$, in which each column is a column vector of length $L$, the dictionary learning process is to find a matrix $D$ that is able to represent the training set with a set of coefficients $X = [x_1, x_2, \ldots, x_n]$ satisfying the specified structural pattern described by the measure $\Omega(x)$ by solving the following optimization problem:

$$\{\hat{D}, \hat{X}\} = \arg\min_{D,X} \{||\Psi Y - DX||^2_F, \Omega(x_i), i = 1, 2, \ldots, n\} \quad (2.1)$$

Herein, the matrix $\Psi$ can be a well-defined analytical transform for incorporating some prior knowledge about the data $Y$, of course, including the special case $\Psi = I$.

Learning simultaneously the dictionary $D$ and the coefficients $X$ corresponds to a matrix factorization problem of $Y$ Tosić and Frossard, 2011. One of popular dictionary learning algorithm is so-called K-SVD, where the alternative iterative strategy is implemented to update $D$ and $X$ until arriving some stop criterion, more specifically; the SVD (the single value decomposition) operation is used to calculate $D$. The standard K-SVD algorithm (corresponding to $\Psi = I$) operates on low-dimensional and typically fixed-size signals, and build a dictionary which is an unconstrained and unstructured matrix. To cope with high-dimensional data with the induced complexity and memory requirements, a structured dictionary is appealing. Rubinstein et al., 2010, developed the sparse K-SVD Rubinstein et al., 2010, behind which the learned dictionary atoms from K-SVD may still share some underlying sparse pattern over a generic dictionary (namely, $\Psi$ is taken as the DCT, Wavelet, Curvelet, and so on). This method provides a parametric model of the learned dictionary, which strikes a good balance among complexity, adaptivity, and performance. More recently, Zhu et al. developed another useful variant on K-SVD, which only requires learning a sparse matrix $D$, and each atom in the learned dictionary is itself a linear combination of atoms in the base dictionary Zhu et al., 2015. Such method improves the efficiency and stability of dictionary learning and provides a new layer of adaptivity to the existing efficient transforms. Recently, Ravishankar and Bresler developed the sparsifying transforming learning method
with the guarantee of stable convergence and low-complexity learning algorithm Ravishankar and Bresler, 2013; Ravishankar and Bresler, 2015. Later on, they also developed the blind compressive sensing model by which the unknown structural model and signal can be reconstructed simultaneously from measurements in combination with the overlapping patch idea Ravishankar and Bresler, 2016, as opposed to other compressive sensing techniques with fixed sparsity model. Finally, it is worthy of remarking that, besides above techniques of dictionary learning, the techniques developed out of machine learning also can be utilized for modeling structures of considered task-driven signals, for instance, the Gaussian process kernel learning Rasmussen and Williams, 2006, the restricted Boltzmann machine Dremeau et al., 2012, and so on.

2.2 Basic Theories

The model-based electromagnetic imaging aims at enhancing the imaging performance by encoding the additional structure of probed objects into the imaging procedure. Two fundamental questions immediately come up: how does the structure information of the image scene affect the imaging performance? What is the ultimate limit on the imaging performance if exploring the structure information into imaging procedure? By now, a body of theories has been established for a few special structure measures like $\ell_1$-norm Duarte and Eldar, 2011; Candès and Tao, 2006; Candès and Tao, 2005; Candès et al., 2004; Candès et al., 2006; Donoho, 2006a; Baraniuk et al., 2010a and nuclear norm Wainwright, 2014; Candès and Recht, 2012; Candès and Plan, 2009; however, two questions above remain open for general structure measures, even for $\ell_1$-norm in some cases. For instance, Candès et al. have built a theory that demonstrates the effect of the TV regularizer on imaging resolution for the partial Fourier or random measurement matrix Candès and Fernandezgranda, 2013; Candès and Fernandez-Granda, 2012. Although the Fourier measurement matrix is widely used for modeling the far-field optical imaging; it is indeed that the measurement matrix is significantly different from the Fourier matrix for most of practical applications of electromagnetic imaging. Nonetheless, we have witnessed that the use of low dimensional prior is capable of profoundly enhancing the imaging
performance and quality almost over the whole electromagnetic spectrum Candès et al., 2004; Massa and Teixeira, 2015; Massa et al., 2015; Potter et al., 2010; Hunt and Smith, 2013; Li et al., 2015; Stevanovic et al., 2016a; Anselmi et al., 2015; Scapaticci et al., 2014; Bevacqua et al., 2015; Winters et al., 2010; Hurtado et al., 2011; Solimene et al., 2012; Soldovieri et al., 2012; Oliveri et al., 2011; Baraniuk and Steeghs, 2007; Cetin and Karl, 2001; Cetin et al., 2004; Huang et al., 2010. This section summarizes several developed theories relevant to the electromagnetic imaging. Comprehensive discussions can be found in well-texted books Elad, 2010 and review papers Elad, 2012; Duarte and Eldar, 2011.

So far, various theories have been established to study the performance of model-based signal representation and recovery; more strictly, the sparse signal recovery, for example, the coherence-based analysis, null-space-based analysis, and RIP-based analysis. Most recently, Tang and Nehorai developed verifiable sufficient conditions and computable performance bounds of \(l_1\)-minimization based sparse recovery by introducing the \(l_\infty\)-norm of errors as the performance criterion Tang and Nehorai, 2015. Nonetheless, we adopt the RIP-based theory for model-based signal recovery, since its bound on the estimation error is tightest over others in general. The RIP implies incoherence and that the linear measurement operator \(A\) approximately preserves the geometry of all sparse vectors. Formally, it defines the restricted isometry constant \(\delta_k\) to be the smallest non-negative constant such that Candès et al., 2006 for all \(k\) sparse vectors \(x\):

\[
(1 - \delta_k)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_k)\|x\|_2^2
\]  

(2.2)

The Candès-Tao-Romberg theorem is stated as following:

**The Candès-Romberg-Tao Theorem (Candès et al., 2006):** Let \(A\) be a measurement operator which satisfies the RIP. Then, for any signal \(x\) and its noisy measurement \(y = Ax + n\) with \(\|n\|_2 \leq \varepsilon\), the solution \(\hat{x}\) to Eq. (2.3) satisfies

\[
\|\hat{x} - x\|_2 \leq C \left( \varepsilon + \frac{\|x_k - x\|_2}{\sqrt{k}} \right)
\]
where $x_k$ denotes the vector of the $k$ largest coefficients in magnitude of $x$. Eq. (2.3) involved above is defined as follows:

$$\min_x \|x\|_1, \quad s.t. \quad \|y - Ax\|_2 \leq \varepsilon$$

(2.3)

This result provides uniform guarantees with optimal error bounds using few non-adaptive samples. The important question is what type of sampling operators has the RIP property, and how large does the number of samples have to be. Fortunately, the literature in compressive sensing has shown that many classes of matrices such as sub-Gaussian, Bernoulli, and so on possess RIP with overwhelming probability when the number of measurements $m$ is nearly linear in the sparsity $k$, especially for the large-scale case Donoho, 2006a. However, it is computationally prohibitive for an arbitrary measurement matrix to show whether satisfies RIP with useful tight bound or not, which remains challenging and open. A tractable strategy alternative to the RIP criteria is the phase transition map calculated via the Monte-Carlo simulation, which has been advocating by Donoho et al. Donoho and Tanner, 2010. Apparently, another interesting benefit from the phase-transition criterion over RIP based analysis is the effect from reconstruction algorithm performance can be taken into account automatically. Since the birth of RIP-based CS theory for sparse analysis, many extensions have been suggested, Li, 2012; Pournaghi and Wu, 2014; Zhao et al., 2013. For instance, the D-RIP copes with sparse synthetic over arbitrary dictionaries. The definition of D-RIP is Tang and Nehorai, 2015:

**Definition 5:** Assuming that $A$ is a given measurement operator and that $D$ is a general sparse synthetic operator. There is a smallest non-negative constant $\delta_D^k$ such that

$$\left(1 - \delta_k^D\right) \|Dx_k^D\|_2^2 \leq \|Ax_k\|_2^2 \leq \left(1 + \delta_k^D\right) \|Dx_k^D\|_2^2$$

(2.4)

holds for all vectors which are $k$-sparse through the transformation $D$. Here, the quantity $x_k^D$ indicates $Dx$ is $k$-sparse.

Note that when $D$ is a normalized tight frame, the D-RIP is automatically reduced into the standard RIP. On basis such D-RIP, the
Candès-Tao-Romberg theorem can be extended as following:

**The Generalized Candès-Tao-Romberg Theorem (Li, 2012):** Assuming that the measurement operator $A$ and the sparse synthetic operator $D$ satisfy the D-RIP with $\delta^D_K < \sqrt{2} - 1$. Then, for any signal $x$ and its noisy measurement $y = Ax + n$ with $\|n\|_2 \leq \varepsilon$, the solution $\hat{x}$ to Eq.(2.5) satisfies:

$$\|Dh\|_2 \leq C_0 \frac{\sigma_k(x)}{\sqrt{k}} + C_1 2 \sqrt{1 + \delta^D_{2k}} \varepsilon; \text{ if } B(y) = \{x, \|Ax - y\|_2 \leq \varepsilon\}$$

$$\|Dh\|_2 \leq C_0 \frac{\sigma_k(x)}{\sqrt{k}} + C_1 2 \sqrt{2k} \lambda; \text{ if } B(y) = \{z, \|A^T (Az - y)\|_\infty \leq \lambda\}$$

And $\|Dh\|_2 \leq C_0 \frac{\sigma_k(x)}{\sqrt{k}} + C_1 2 \sqrt{2k} \lambda \| (D^T D)^{-1} \|_\infty; \text{ if } B(y) = \{z, \|A^T (Az - y)\|_\infty \leq \lambda\}$

Eq. (2.5) used above theorem is defined as following constrained convex optimization problem:

$$\hat{x} = \arg \min_z \|Dz\|_1 \quad \text{s.t.} \quad z \in B(y) \quad (2.5)$$

where $B(y)$ includes three typical cases: (a) Noise-free case $B(y) = \{z, \phi z = y\}$, (b) Dantizig selector case, $B(y) = \{z, \|Az - y\|_2 \leq \varepsilon\}$, and (c) Robust optimization case, $B(y) = \{z, \|A^T (Az - y)\|_\infty \leq \lambda\}$.

The CS theory paves a new paradigm for task-driven signal processing by exploiting structure property in the procedure of data acquisition and processing, which also beats heavily problems thought to be impossible in the past. There are many powerful extensions on CS theory, which are relevant to the electromagnetic imaging, as summarized in Table 2.1. The blind deconvolution basically aims at estimating both the signal $x$ and the unknown convolution kernel function $h$ from the noisy measurement $y$, which is challenging when $x$ is of non-minimum phase. However, many encouraging theoretical and experimental results could be achieved if the vector $x$ is sparse in the aid of structural signal processing. This type of problems has found their valuable applications in many areas of radar imaging Li, 2014. For example, the ground-penetrating radar imaging consists of estimating the spark-like reflectivity sequence, where the transmitting radar wavelet (corresponding to $h$) is typically unknown in practice and required to be estimated simultaneously Li, 2014. The sparsity-promoted blind deconvolution problem also can be applicable for more general problem of radar imaging where the radar signal suffers from the ambiguous background medium, for instance, the
space-borne radar imaging affected by the ionosphere, the through-wall radar imaging with ambiguous wall’s parameters. Phase retrieval is another typical problem for the intensity-based electromagnetic imaging with relatively high operational frequencies, such as millimeter, THz, and so on Li et al., 2008; Li et al., 2009. Even for the microwave regimes, phase retrieval is also helpful in saving hardware cost and system complexity. The multi-task Ji et al., 2009 (or the multi measurements vectors, MMV, is called in some literatures Cotter et al., 2005) CS aims at treating a series of measurements \( \{ y_i \} \), where the series of sparse vectors \( \{ x_i \} \) share certain joint sparse property as discussed below. Such kind of problem is widely involved in the electromagnetic imaging problem, typically, the multi-dimensional radar imaging where the reflectivity of objects are different for different view angles, operational frequencies, polarization states, etc.. Most recently, Thrampoulidis et al. demonstrated that the nonlinear measurements \( y_i = g_i(\langle a_i, x \rangle) \) \((i = 1, 2, \ldots, M)\) for the structure signal is equivalent to the linear problem of \( y_i = \mu \langle a_i, x \rangle + \sigma z_i \) (where \( \mu \) and \( \sigma \) are known parameters as a function of \( g_i \)) by solving the robust LASSO problem \( \min_{x} [\| y - Ax \|_2 + \gamma f(x)] \), which provides a new unified perspective on several nonlinear CS methods including one-bit CS and phase retrieval Thrampoulidis et al., 2015.
### 2.2. Basic Theories

Table 2.1: Some representative CS extensions for electromagnetic imaging.

<table>
<thead>
<tr>
<th>Term</th>
<th>Problem Statement</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-bit CS (Single-index Measurement)</td>
<td>Find the sparse vector $x$ from $y = \text{sgn}(Ax + n)$ where $y \in [0, 1]$ $\text{sgn}(.)$ is a binary operator</td>
<td>Hardle et al., 1991; Hristache et al., 2001; Gopi et al., 2013; Plan and Vershynin, 2011; Plan and Vershynin, 2012</td>
</tr>
<tr>
<td>Low-matrix Completion</td>
<td>Find low-rank matrix $X$ from $y = Ax + n$</td>
<td>Tosic and Frossard, 2011; Rubinstein et al., 2010; Recht et al., 2010</td>
</tr>
<tr>
<td>Phase Retrieval</td>
<td>Find the sparse vector $x$ from $y = Ax + n$</td>
<td>Bahmani and Romberg, 2015</td>
</tr>
<tr>
<td>Blind deconvolution</td>
<td>Find $x$ and $h$, s.t. $y = x \ast h$, Here $\ast$ is convolution operator.</td>
<td>Ahmed et al., 2012</td>
</tr>
</tbody>
</table>
Reconstruction Algorithms

The model-based electromagnetic imaging heavily relies on the implementation of efficient reconstruction algorithm. This problem is typically in lack of closed-form solution and usually suffers from some uncertainties caused by sensors’ accuracy and ambiguous system parameters, so we would like to treat it in the systematic framework of convex optimization. Convex optimization has been studied over a century and nowadays has become a routine job in many disciplines Boyd and Vandenberghe, 2004; Darzentas, 1984. However, convex optimization by itself is being under a tremendous pressure in the big-data era, where the encountered problem is of very large scale because it is desirable to accommodate extremely massive measurements with the increasing demands of retrieving very detailed information of undergoing samples. Three basic considerations for treating the very large-scale and high-resolution electromagnetic imaging in the framework of convex optimization are:

a. The measurement operator $A$, which approximates the real physical interaction between electromagnetic wavefields and probed objects, is built such that the optimal trade-off between the imaging quality and the computational cost can be achieved.
b The regularizer $\Omega$, by which the low-dimensional prior on probed object is incorporated into the imaging procedure for the purpose of enhancing the imaging quality, is selected to support the development of efficient reconstruction algorithms. Specially, it is rather pleasant thing that the measure has some structure properties like smoothness, strong convexity, proximal tractability, and decomposability.

c It is of central importance to develop reconstruction algorithms with faster convergence rate and lower computational complexity. Moreover, it is also of significance to algorithms that are both rich enough to capture the structures of undergoing problem, and are scalable enough to support the realization in a parallelized or fully distributed fashion.

The point (a) listed above is relevant to the physical mechanism of electromagnetic imaging, and is left for discussion in next section. This section is focused on the points (b) and (c). Although a great amount of algorithms have been developed to efficiently and reliably solve convex optimization problems (for example see Boyd and Vandenberghe, 2004 for references herein), we would like to restrict ourselves into the first-order methods founded by Nesterov Nesterov, 2013 and Nemirovskii Darzentas, 1984; Nemirovski et al., 2009 for handling model-based electromagnetic imaging due to following considerations. First, as implied by the name, first-order methods only involve the operation of full or stochastic or partial gradient, which has apparently low computational cost as opposed to other techniques like Gaussian-Newton methods requiring the operation of Hessian matrix with really expensive computational cost. Second, despite their solutions with low- or medium- accuracy, first order methods can handle the non-smooth problem, a typical situation in the sparsity-driven imaging, by making use of the proximal mapping technique. Third, it turns out that first-order methods are quite robust to use approximations of their optimization primitives, i.e., randomization techniques. Key ideas behind randomization techniques include random partial updates of optimization variables, replacing the deterministic gradient and proximal calculations with cheap statistical
estimators, and speeding up basic linear algebra routines via randomization Cevher et al., 2014; Bahlak et al., 2015. Finally, first order methods naturally provide a flexible framework to distribute optimization tasks and perform computations in the parallelized or distributed manner Cevher et al., 2014. For above considerations, in what following we will review three important members of first order methods.

### 3.1 First-order Proximal Methods

Recalling Eqs. (1.1), the model-based electromagnetic imaging consists of solving following unconstrained optimization problem, i.e.,

$$
\min_{x} [f(x) + \Omega(x)]
$$

(3.1)

Herein, the first term $f(x)$ defines the data fidelity which is usually smooth and strongly convexity, e.g., $f(x) = \frac{1}{2} \| Ax - y \|_2^2$. However, the penalty term $\Omega(x)$ is usually non-smooth, even more, non-convex. In the black-box model, the non-smoothness dramatically deteriorates the rate of convergence of first-order methods by a factor of $O(\sqrt{k})$. A natural idea of dealing with non-smooth $\Omega(x)$ is to explore smoothing techniques including the Nesterov’s smoothing technique Nesterov, 2005b, and the smoothing relaxation via small perturbation Figueiredo and Nowak, 2003, Electronic Warfare and Radar Systems Engineering Handbook 1997. On the other hand, the penalty $\Omega(x)$ has some attractive properties, which admits, if exploited, that the non-smooth problem of Eq. (3.1) can be solved nearly as efficiently as smooth problems. First-order proximal methods play this role, which is centered on the concept of proximal operator. The definition of proximal operator was introduced for the quantity of real-valued vector by Moreau (1965) Bahlak et al., 2015, whose slight extension reads:

**Definition 6:** For a real-valued function (either smooth or non-smooth) $f(X)$, its proximal operator is defined as:

$$
\text{Prox}_f(Y) = \arg \min_{X} \left[ f(X) + \frac{1}{2} \| Y - X \|_2^2 \right]
$$

(3.2)

Here $\| \cdot \|$ denotes the Frobenius norm if $X$ is a matrix, and the Euclidean norm if $X$ is a vector.
3.1. First-order Proximal Methods

For every point $X$ in high dimensional space, $\text{Prox}_f(X)$ can be interpreted as the minimizer of $f$ in the neighborhood of $X$; moreover, the solution to this minimizer is always unique Komodakis and Pesquet, 2014. Table 3.1 summarizes some useful proximal operators with closed-form expressions encountered in structure electromagnetic imaging. Here we would like to emphasize that the structure measure $\omega(x) = \sum g \alpha_g \|x_g\|_1 + \sum g \omega_g \|x_g\|_2$, where $G$ denotes the total number of groups, $\{\alpha_g, g = 1, 2, \ldots, G\}$ and $\{\omega_g, g = 1, 2, \ldots, G\}$ are known weighting factors, is useful in modeling variables partitioned into a few groups while elements in each group are sparse.

| $\Omega(x)$ | $\text{Prox}_\Omega$ | Ref.
|---|---|---
| $\|x\|_1$, $x$ is real | $(|x_{0,i}| - \mu)\text{sgn}(x_{0,i})$ if $|x_{0,i}| \geq \mu$; 0, else. | Tipping, 2001
| $\|x\|_1$, $x$ is complex | $(|x_{0,i}| - \mu)A_g(x_{0,i})$ if $|x_{0,i}| \geq \mu$; 0, else. | Goldstein et al., 2015
| $\sum g \|x_g\|_2$ Non-overlapping | $(\|x_g\|_2 - \mu)\frac{x_g}{\|x_g\|_2}$ if $\|x_g\|_2 \geq \mu$; 0, else. | Tipping, 2001
| $\sum g \alpha_g \|x_g\|_1 + \sum g \omega_g \|x_g\|_2$ Non-overlapping | $\frac{\Delta_g - \omega_g}{\Delta_g}(|x_{g,0,i}| - \alpha_g)\text{sgn}(x_{g,0})$ if $|x_{g,0,i}| \geq \alpha_g$; 0, else. $\Delta_g = \sqrt{\|(x_{g,0} - \alpha_g)_+\|_2^2 + \|(x_{g,0} - \alpha_g)_-\|_2^2}$ | This work

Table 3.1: Some closed-form solutions of $\text{Prox}_{\mu\Omega}(p(y))$.

It should be highlighted that various proximal operators of popular structure measures mentioned previously have no closed-form solution, and thus numerical strategies are needed Goldstein et al., 2015. A representative example is $\Omega_s(x) = \Omega(D(x))$, where $D$ is a known sparsifying transform. This type of problems can be tackled by using
the splitting-operator technique as well as (augmented) Lagrangian technique Boyd et al., 2011, behind which the key idea is to introduce an auxiliary quantity $u = D(x)$, and then the alternative direction method is employed to solve the resulting optimization problem over the pair of prime-dual variables $(x, u)$. An alternative (actually, equivalent in some cases) technique is the use of the Fenchel or Legendre transform of $\Omega_s(x)$ to form a min-max saddle optimization problem (i.e., primal-dual problem) Komodakis and Pesquet, 2014, which also can be efficiently solved by performing the alternative iterative scheme. Nonetheless, it is not a trivial issue to obtain numerically the solution to the proximal operator of $\Omega_s(x)$ for (very) large-scale problems, especially for those transforms $D$ lacking of fast forward and/or inverse transform algorithms.

Now, we turn to the discussion of first-order proximal methods. A fundamental fact central to first-order methods is that $f(x)$ is assumed to be $L$-Lipschitz smooth in the Euclidean norm Nesterov, 2013, i.e., $\|\nabla f(x) - \nabla f(y)\| \leq L \|x - y\|_2$, for any $x, y \in \text{dom}(f)$; equivalently, the upper bound inequality of $f(x) \leq f(y) + \langle \nabla f(y), x - y \rangle + \frac{L}{2} \|y - x\|_2^2$ holds for any $x, y \in \text{dom}(f)$. First-order proximal methods can be described as follows Bahlak et al., 2015: starting from an initial guess $x_0$, and then iterates the following equation:

$$x_{(k+1)} = \min_x \left[ f(x_k) + \langle \nabla f(x_k), x - x_k \rangle + \frac{L}{2} \|x - x_k\|_2^2 + \Omega(x) \right]$$

$$= \min_x \left[ \frac{1}{2} \|x - p(x_k)\|_2^2 + L^{-1} \Omega(x) \right]$$

$$\equiv \text{Prox}_{L^{-1}\Omega}(p(x_k)) \quad (3.3)$$

where $p(x_k) = x_k - L^{-1}\nabla f(x_k)$ characterizes exactly the first-order iterative solution of $f(x)$ in the proximity of $x_k$. Taking $f(x) = \frac{1}{2} \|Ax - b\|_2^2$ as example, $p(x_k) = y - L^{-1}A'(Ax_k - b)$ represents the standard back-projection imaging operation Bahlak et al., 2015 (which becomes more clear for $p(0) = L^{-1}A'b$), while $\min_x [\frac{L}{2} \|x - p(x_k)\|_2^2 + \Omega(x)]$ corresponds to the imaging processing where the low-dimensional prior is taken into account. In language of electromagnetic imaging, the iterative solution of Eq. (3.3) is the closed loop consisting of back-projection and structural image processing, as illustrated in Figure 3.1, as opposed
to traditional open-loop processing strategies in the area of microwave imaging Zhdanov, 2009. In this sense, the model-based electromagnetic imaging is a natural bridge filling the gap between the electromagnetic imaging and image process and pattern recognition, which is exactly basic purpose of computational electromagnetic imaging.

Regarding the determination of $L$ in Eq. (3.3), taking $f(x) = \frac{1}{2}\|Ax - b\|_2^2$ as example, the most suggested strategy is the linear search. For general $f(x)$, it may be computationally prohibitive for large-scale problems. The convergence rate will be very slow for relatively small value of $L$. However, a big value cannot guarantee the stable convergence, which comes from the fact that the bound inequality $f(x) \leq f(y) + \langle \nabla f(y), x - y \rangle + \frac{L}{2}\|y - x\|_2^2$ is broken. As pointed out previously, above first-order methods with constant step-size $\frac{1}{L}$ have very low convergence rate of $O\left(\frac{L}{K}\right)$ Bahlak et al., 2015. To beat it, accelerated first-order methods with the rate of super linear convergence have been suggested, e.g. the Nesterov’s accelerated gradient method Maleki et al., 2013—see a detailed discussion in Bahlak et al., 2015. The original Nesterov’s method was developed for smooth problems, and has been remarkably generalized for general situations by many researchers Nesterov, 2005b; Beck and Teboulle, 2009; Tseng, 2008; Becker et al., 2009; Su et al., 2015. For example, its variant known as FISAT was suggested for the non-smooth problem Eq. (3.1) with $\Omega(x) = \|x\|_1$. We would like to refer to Allen-Zhu and Orecchia for its relation to mirror descent AllenZhu and Orecchia, 2014, and to Tseng for a unified analysis of these ideas Tseng, 2008. Kim and Fessler established an analysis on the worst bound of Nesterov’s method, and further proposed two optimized first-order algorithms that achieves a convergence bound that
is two times smaller than for Nesterov’s fast gradient methods Kim and Fessler, 2016. Su et al. developed a differential-equation interpretation of Nesterov’ methods with a new restarting scheme, which overcomes the challenging issue of estimating the step-size and leads to linear convergence rate whenever the objective is strongly convex Su et al., 2015.

### 3.2 First-order Primal-dual Methods

Primal-dual methods are another popular approaches with the solid guarantee of convergence rate, which holds the promising for addressing very large-scale model-based electromagnetic imaging due to several unique advantages including easy-implementation, low computational complexity, supporting parallel and distributed realization, incorporating the strength of proximal methods (e.g., not only smooth but also non-smooth structural measures), etc Komodakis and Pesquet, 2014. As their name implies, primal-dual methods proceed by concurrently solving a primal problem (corresponding to the original optimization task) as well as a dual formulation of this problem. Recently, an interesting observation on primal-dual methods draw by Zhang et al. is that the dual solution is also sparse or at least approximately sparse if the primal solution is sparse with high prediction accuracy Maleki et al., 2013. More interestingly, primal-dual techniques are able to achieve what is known as full splitting in the optimization literature, meaning that each of the operators involved in the problem (i.e., not only the gradient or proximity operators but also the involved linear operators) is used separately Boyd et al., 2011. Primal-dual methods are usually derived in the context of the method of Lagrangian multiplier, which are primarily applicable to the convex optimization problem where the strong duality holds Boyd et al., 2011. A more general framework of primal-dual methods without the requirement of strong duality is summarized below:
Statement on general primal-dual framework: For the constrained convex optimization problem

\[ \min_x \Omega(x), \text{s.t.}, h_i(x) \leq 0, i = 1, 2, 3..., n \]  

(3.4)

Then, the unconstrained primal-dual problems of Eq. (3.4) read:

(PP) \[ \min_x p(x) \]  
(DP) \[ \min_\lambda d(\lambda) \]

where \( \lambda \) is the (Lagrangian) dual variable, the primal function \( p(x) \) and the dual function \( d(\lambda) \) are \( p(x) = \max_{\lambda \geq 0} [\Omega(x) + \sum_{i=1}^{p} \lambda_i h_i(x)] \) and \( d(\lambda) = \min_x [\Omega(x) + \sum_{i=1}^{p} \lambda_i h_i(x)] \), respectively. For any \( x \) and \( \lambda \), there is the weak-duality or duality gap \( f(x) = p(x) \geq d(\lambda) \), moreover, the strong duality implies that \( p = d \). The primal function \( p(x) \) and dual function \( d(\lambda) \) are usually non-smooth, which can be smoothed via the Nesterov’s smooth technique as

\[ p_\beta(x) = \max_{\lambda \geq 0} [\Omega(x) + \sum_{i=1}^{p} \lambda_i h_i(x) - \frac{\beta}{2} \| \lambda \|_2^2] \]  

(3.5)

and

\[ d_\mu(\lambda) = \min_x [\Omega(x) + \sum_{i=1}^{p} \lambda_i h_i(x) + \frac{\mu}{2} \sum_{i=1}^{p} h_i^2(x)] \]  

(3.6)

respectively, where \( \beta \) and \( \mu \) are referred to as smooth factors. Other smoothing techniques are also applicable for \( p_\beta(x) \) and \( d_\mu(\lambda) \) Maleki et al., 2013. Then, the pair of smoothed primal-dual problems are:

(PP\( _\beta \)) \[ \min_x p_\beta(x) \]  
(DP\( _\mu \)) \[ \min_\lambda d_\mu(\lambda) \].

With above notations, primal-dual methods developed by now can be roughly categorized into three groups:

### 3.2.1 Lagrangian multiplier based primal-dual methods

This kind of PD methods is aimed at dealing with one of non-smooth problems DP and PP. Actually, it can be immediately observed that DP and PP are equivalent to each other. There are two drawbacks standing out others that, for instance for solving PD, (a) it is not a trivial issue to find the update step-size of \( \lambda \) for solving \( p(x) \), (b) pretty slow
convergence rate with the order of $O\left(\frac{1}{\sqrt{k}}\right)$ Boyd et al., 2011, where $k$ is the iterative number. For these reasons, the Lagrangian multiplier based PD methods are usually not suitable for solving large-scale model-based electromagnetic imaging problems.

### 3.2.2 Augmented Lagrangian multiplier based primal-dual methods

This type of PD methods are designed to solve the smoothed problem $\text{DP}_\mu$. For example, considering the case of $h_i(x) = a_i^T x$ or equivalently,

$$d_\mu(\lambda) = \min_x [\Omega(x) + \lambda^T (Ax - b) + \frac{\mu}{2} \|Ax - b\|_2^2] \quad (3.7)$$

Here, the smoothing factor $\mu$ plays the role of enhancing the strong convexity of objective function by the Lipschitz smooth constant of $L_f + \mu \|A\|_2^2$, which consequently leads to improved algorithmic convergence performance. Classical alternative iterative technique solving $\text{DP}_\mu$ produces a sequence of $\{(x_k, \lambda_k), k \geq 0\}$ starting from $(x_0, \lambda_0)$ as Boyd et al., 2011

$$\begin{cases}
x_{k+1} = \arg \min_x \left[ \lambda_k^T (Ax - b) + \frac{\mu}{2} \|Ax - b\|_2^2 + \Omega(x) \right] \\
\lambda_{k+1} = \lambda_k + \mu (Ax_{k+1} - b)
\end{cases} \quad (3.8)$$

or

$$\begin{cases}
x_{k+1} = \arg \min_x \left[ \frac{\mu}{2} \|Ax - b + u^k\|_2^2 + \Omega(x) \right] \\
u_{k+1} = u_k + (Ax_{k+1} - b)
\end{cases} \quad (3.9)$$

Eq. (3.9) is the scaled version of Eq. (3.8), where $u = \frac{\lambda}{\mu}$. Besides Eq. (3.8) and (3.9), many variants on it have also been developed to solve $\text{DP}_\mu$. It has been rigorously proved that the method (3.8) and (3.9) and its many variants has the convergence rate of $O\left(\frac{1}{k}\right)$ under mild condition; however, its accelerated version with the convergence rate of $O\left(\frac{1}{k^2}\right)$ can be developed using Nesterov’s accelerated technique Maleki et al., 2013.

Despite the moderate solution accuracy compared to that obtained by the interior point programming and other methods, the method (3.9) and its variants have three-fold interesting properties making them very
3.2. First-order Primal-dual Methods

popular in solving large-scale problems. The first property is that they have no call for the computation of matrix inversion during the iteration procedure. Second, more advanced optimization schemes with lower computational cost and faster convergence rate can be immediately developed by exploiting the properties of input function in these PD methods. For instance, the decomposability of $\Omega(x)$ supports immediately the realization of PD algorithm in the parallel and distributed manner. The tractability of proximal operator of $\Omega(x)$ can enhance computation efficiency by marrying PD methods with proximal methods, by which the non-smooth optimization problem can be solved as nearly efficiently as smooth problem. For above considerations, primal-dual methods lead to algorithms that are easily parallelizable, which is nowadays becoming increasingly important for efficiently handling high-dimensional problems. Nonetheless, we would like point out that the operation of matrix inversion, in spite of that calculated in the off-line manner, is involved for primal-dual methods, which is really expensive for very-large-scale electromagnetic imaging problem, even computationally prohibitive in some cases. Two aspects can be considered for addressing it. The linearization technique can be explored to solved the first line of Eq. (3.8) and (3.9), i.e.,

$$x^{k+1} \approx \arg \min_x \left[ \frac{L_A}{2} \| x - \frac{1}{L_A} A' (\lambda_k + \mu (Ax_k - b)) \|_2^2 + \Omega(x) \right]$$

$$= \text{Prox}_{\mu^{-1} \Omega(x)} \left( \frac{A' (\lambda_k + \mu (Ax_k - b))}{L_A} \right)$$

(3.10)

for Eq. (3.8) or

$$x^{k+1} \approx \arg \min_x \left[ \frac{L_A}{2} \| x - \frac{\mu}{L_A} A' (Ax_k - b + u_k) \|_2^2 + \Omega(x) \right]$$

$$= \text{Prox}_{\mu^{-1} \Omega(x)} \left( \frac{\mu}{L_A} A' (Ax_k - b + u_k) \right)$$

(3.11)

for Eq. (3.9), where $L_A = \mu \| A \|_2^2$. If the proximal operator of $\Omega(x)$ is analytically tractable, the linearized ADMM, known as the preconditioned ADMM Boyd et al., 2011, is of low-complexity, completely surpass the operation of expensive matrix inversion, as desired by solving the large-scale convex optimization problem. Third, many applications of
practical electromagnetic imaging support the efficient computation of the first line of Eq. (3.8) and (3.9) with attractive properties, such as the imaging model $A$ in the far-field imaging application can be justified as a Fourier-transform matrix with acceptable accuracy, and $A$ is also can be regarded as a convolution operator. As it turns out, in doing so they are able to exploit more efficiently specific properties of the measurement operator $A$ and structural measure $Ω(x)$, thus offering in many cases important computational benefits.

### 3.2.3 Excessive gap based primal-dual methods

This kind of methods are new members of primal-dual methods, which are aimed at solving jointly both PP$_β$ and DP$_μ$ such that the excessive gap of $G_{μβ}(w) = d_μ(λ) − p_β(x) ≥ 0$ holds, where $w = [x^T, λ^T]^T$. The concept of excessive gap was introduced by Nesterov, 2005a, which comes from a fundamental fact the non-smooth PP and DP problems have always the duality gap $p(x) ≥ d(λ)$; however $p(x) ≥ p_β(x)$ and $d_μ(λ) ≥ d(λ)$, as a consequence, it can be faithfully excepted that a gap (termed as the excessive gap) of $d_μ(λ) ≥ p_β(x)$ exists. As a result, once a solution is calculated by the excessive gap based primal-dual methods, this solution is the optimal solution to (PP) and (DP) in the limit of $μ → 0$ and $β → 0$. Moreover, it can be proven that $0 ≤ \max\{p(x) − p^∗, p^∗ − d(λ)\} ≤ p(x) − d(λ) ≤ μD_1 + βD_2$, where $D_1 = \max_x p(x)$ and $D_2 = \max_λ d(λ)$ Nesterov, 2005a. Additionally, assuming that $λ^∗(x, β) = \frac{1}{β}(A'x − b)$ is the optimal solutions to $p_β(x)$ at a given $x$, while $x^∗(λ, μ)$ denotes the optimal solutions to $d_μ(x)$ at a given $λ$. For a given pair of solution $(x, λ)$ satisfying the excessive gap, then the solution $(x^+, λ^+)$. produced by the P2D1 or P1D2 algorithms satisfies the excessive gap as well under mild condition Trandinh and Cevher, 2015. Note that all parameters including $β, τ$ and $μ$ involved above can be analytically calculated in the iterative process, we would like to refer the readers for details in Trandinh and Cevher, 2015. More interestingly, Tran-Dinh and Cevher provided optimal convergence rates on the primal objective residual $O(\frac{1}{kα}) (α=1,2)$ as well as the primal feasibility gap $O(\frac{1}{kα}) (α=1,2)$, separately.
3.3 Randomized First-order Methods

Most of practical high-resolution electromagnetic imaging problems are
of very large-scale, which are typically of massive measurements and
high-dimensional variables, as highlighted frequently in previous sections. It is interesting that massive measurements are usually redundant and statistically follow an independent identical distribution. More specifically, each measurement reflects the behavior of the whole data. In terms of computational efficiency, this property is important since it supports not only the randomized dimensional reduction, but also, more importantly, the reasonable approximation of the gradient involved in first-order methods with randomized gradient instead of full gradient. On the other hand, first order methods are quite robust to the approximation of gradient computation. Therefore, randomized first-order methods exhibit significant acceleration over their deterministic counterparts since they can generate a good quality solution with high probability by inspecting only a negligibly small fraction of the data or variables. Moreover, since the computational primitives of such methods are inherently approximate, we can often obtain near linear speed-ups with a large number of processors.

3.3.1 Randomized Coordinate Descent (RCD) Methods

Randomized coordinate descent methods have a long history in optimization and are related to many classical methods in diverse fields, for example, the alternative iterative method in the context of convex optimization Boyd et al., 2011, the randomized Gauss–Seidel method in the area of classical linear algebra, the ordered subsets method widely used in the area of medical imaging Bai et al., 2002; Ahn and Fessler, 2003, and the randomized simultaneous perturbation sequentially algorithm (SPSA) in the area of signal processing Frey, 2004; Li et al., 2013. The central idea of randomized coordinate descent methods is that each iteration only one or several variables are randomly picked, and updated to improve the objective function as done by first-order methods. These methods require only one-vector operation at lower computational cost, which differs from full-gradient iterative method requiring calculating
Reconstruction Algorithms

full gradient each iteration. Borrowing from results of ordered subsets, the randomized coordinate descent method can be demonstrated to be with the accelerated convergence rate at the iterative early stage Ahn and Fessler, 2003. Another benefit made by the randomized coordinate descent method is they can be flexible in treating non-linear problem with the objective function whose gradient is not easy to be calculated Li et al., 2013. A key design consideration across all coordinate descent methods is the choice of the coordinate at each iteration. For the deterministic coordinate descent method, the coordinate with largest direction sensitivity is picked after calculating full gradient Ahn and Fessler, 2003. Finding the best coordinate to update, the maximum of the gradient element’s magnitudes, can require a computational effort as high as the gradient calculation itself. However, the method’s convergence is provably slower than the gradient method due to the basic relationship. The randomized coordinate descent method picks the coordinate(s) in a random manner. Nesterov gave a theoretical prove that the randomized coordinate descent method has the theoretical guarantee of stable convergent property equivalent statistically to full-gradient based optimization method Nesterov, 2010. We refer to Richtarik and Takac Richtarik and Takac, 2014 and Cevher Cevher et al., 2014 for discussions of randomized coordinate descent methods in a distributed setting. More recently, accelerated and composite versions of coordinate descent methods have also been explored Richtarik and Takac, 2014; Fercoq and Richtárik, 2013; Fercoq et al., 2014.

3.3.2 Stochastic Gradient (SG) Methods

Stochastic gradient methods are related to many classical methods, for instance, the celebrated Robbins-Monro algorithm Robbins and Monro, 1985, randomized Kaczmarz Needell et al., 2013; Nemirovski et al., 2009 (one of its variants is the algebraic reconstruction technique (ART) with random selection of data in the area of tomography), online learning Slavakis et al., 2014, and so on. The stochastic gradient method is exactly dedicated for treating decomposable objective function, i.e., $f(x) = \sum_{j=1}^{J} f_j(x) + \Omega(x)$, where each $f_j(x)$ measures the data misfit for the whole variable, and $\Omega(x)$ denotes the structural regularizer. In contrast
to RCD methods, which updates a single coordinate at a time with its exact gradient, SG methods update all coordinates simultaneously but use approximate gradients of partial objective function. Specifically, the gradient at each iteration is approximated by the gradient of one randomly selected measurement $f_j(x)$. A crucial design for stochastic methods is the selection of the data points $f_j$ at each iteration. By now, several typical selection methods have been proposed Qu et al., 2015, i.e., serial uniform sampling, serial optimal sampling (importance sampling), $\tau$-nice sampling, distributed sampling, and so on. Similar to deterministic first-order methods, another central issue to SG methods is the choice of update step size. Indeed, while SG methods have historically been notoriously hard to tune, recent results show that using large step-sizes and weighted averaging of the iterate allows us to achieve optimal convergence rates while being robust to the setting of the step-size and the modeling assumptions Nemirovski et al., 2009. In particular, an averaged SG iteration with a constant step-size achieves an $O(\frac{1}{k})$ convergence rate even without strong convexity under joint self-concordance-like and Lipschitz gradient assumptions. More interestingly, Qu et al. proposed a novel primal-dual stochastic gradient method named as Quartz for arbitrary selection of measurement and adaptive determination of update step size Qu et al., 2015.

Before closing this section, we would like to emphasize that another useful case is that the structure measure is the composition of multiple different individual structural measures. Apparently, above techniques can be exploited for their applicable condition in a straightforward manner after smoothing all possible non-smooth sub-function. Very recently, Yu introduced the proximal average technique for proximal gradient methods (PA-APG) Yu, 2013. Instead of directly solving the proximal step associated with a composite regularizer, it averages the solutions from the proximal problems for each regularizer. It is shown that this approximation is strictly better than that of Nesterov’s smoothing, while enjoying the same per-iteration time complexity and convergence rate. Later on, Zhong and Kwok extended Yu’ method to develop the accelerated stochastic gradient method for composite structural regularizes Zhong and Kwok, 2014.
In the Introduction section, we have pointed out that the electromagnetic imaging can be usually formulated as solving the linear inverse problem, \( y = Ax + n \), where \( y \) indicates the vectorized measurements corrupted with additive noise \( n \), \( x \) denotes the reflectivity of imaged scene to be retrieved. As we frequently have argued previously, such problems usually suffer from several important challenges in practice, mainly due to the ill-posedness arising from the limited number of measurements with respect to that of unknowns to be retrieved, and the heavy computational burden especially for the large-scale imaging scenarios. We have emphasized in Chapter 2 that the low-dimensional representation of underlying scenes can be leveraged to resolve the ill-posedness issue, even for the limited number of measurements, which can be theoretically guaranteed, for instance, by the theory of compressive sensing and its variants. As for the algorithms in efficiently solving the large-scale electromagnetic imaging problems, we have discussed several representative algorithms with relatively low computation complexity, among which the first-order iterative algorithms have been highlighted in Chapter 3. Here, we will focus on the specific formulations of above inverse problems and discuss their applications in practical electromag-
4.1 Electromagnetic Imaging Model

With reference to Figure 4.1, a transmitter and a receiver are located at \( r_t \) and \( r_s \), respectively, while the investigation domain denoted by \( V \) is occupied by the probed object. When the object is illuminated by the electromagnetic wavefield, an induced current denoted by \( J \) will be generated throughout the object. In terms of the electrical integral equation, the electrical field at \( r_s \) radiated from the current \( J \) reads

\[
E(r_s; \omega, r_t) = i\omega \mu_0 \int_V \frac{d r'}{\mathcal{V}} G(r_s, r'; \omega) \cdot J(r'; \omega, r_t)
\]

where \( \omega \) is working angular frequency, \( \mu_0 \) is the free-space permeability, and \( G(r_s, r'; \omega) \) is the dyadic Green’s function in vacuum \( \text{DeVaney, 2012; Pastorino, 2010, i.e.,} \)

\[
G(r, r'; \omega) = \left( I + \frac{1}{k_0^2} \nabla \nabla \right) g(r, r'; \omega)
\]

\[
e^{ik_0 R} \frac{3}{4\pi R} \left( 3 \frac{R \otimes R}{R^2} - I \right) \left( \frac{1}{k_0^2 R^2} - \frac{i}{k_0 R} \right) + \left( I - \frac{R \otimes R}{R^2} \right)
\]

where \( g(r, r'; \omega) \) is the scalar three-dimensional Green’s function in free space, \( k_0 = \frac{\omega}{c} \) denotes operational wavenumber. Introducing formally the notation of \( \overline{\sigma}(r'; \omega, r_t) \) such that \( J(r'; \omega, r_t) = \overline{\sigma}(r'; \omega, r_t) E_{in}(r'; \omega, r_t) \), where \( \overline{\sigma} \) is referred to as the generalized reflectivity of the probed object, and \( E_{in} \) is the incident wavefield emerging from transmitter in the absence of objects. By stacking all measurements and vectoring all unknowns \( \overline{\sigma}(r'; \omega, r_t) \), the electromagnetic imaging problems can be formulated in the form of Eq. (1.1). Then, the basic purpose of electromagnetic imaging is to reconstruct the distribution of \( \overline{\sigma}(r'; \omega, r_t) \) from processing the data generated by using a sequence of transmitters and receivers at different locations. The resulting linear inverse problem
Figure 4.1: The illustrate map for Eq. (4.1)

has been theoretically and practically demonstrated to be seriously ill-posed owing to the existence of so-called non-radiation currents related to multi-scattering effects Devaney, 2012.

Note that the inverse problem of Eq. (4.1) is nearly intractable since the number of unknowns, as implied by the dependence of $\sigma(r'; \omega, r_t)$ on frequencies and transmitters’ locations, is overwhelming over that of measurements. To overcome this issue, some approximations are needed. A well-known case is the Born approximation for weak objects, i.e., $\sigma(r'; \omega, r_t)$ is assumed to be free of frequencies and transmitters’ location. Nonetheless, the resulting problem remains ill-posed, especially for the purpose of high-resolution imaging, as argued previously. To relieve this issue, some prior knowledge on $\sigma$ should be exploited to enhance the imaging performance. In this sense, there is a natural bridge between the electromagnetic imaging and the structure signal processing. From the standpoint of computational efficiency, the matrix constituted by $G(r, r'; \omega)$ has almost no obvious structure to be exploited. As emphasized previously, the model-based electromagnetic imaging needs the implementation of time-consuming iterative reconstruction algorithms. The computational burden is mostly taken on the operation of matrix-vector multiplication, therefore how to fasten the matrix-vector multiplication is appealing especially for large-scale problems. Therefore, it is necessary to approximate $G(r, r'; \omega)$ at the
acceptable sacrifice of solution accuracy. It can be justified that when
the sensor is far away from the object, i.e., \( R = |r - r'| = r - \hat{r}r' \), the
typical situation for the purpose of electromagnetic remote sensing, the
dyadic Green’s function can be approximated as Pastorino, 2010

\[
G(r, r'; \omega) = (I - \hat{k}_s \hat{k}_s) \frac{e^{ikr}}{4\pi r} e^{-i\hat{k}_s r'}
\] (4.3)

Now, the measurements are related to the object by the Fourier
transform like matrix, by which each matrix-vector multiplicative can
be efficiently calculated by the fast Fourier transform.

### 4.2 Compressive Imaging System

In this section, we report recent progress on low-dimensional-model-
based electromagnetic imaging in diverse application areas. Note that, as
surveyed below, the low-dimensional models used in the electromagnetic
imaging are simple sparse or compressible by now. We expect to discover
and incorporate more realistic and richer low-dimensional models in the
electromagnetic imaging in the near future.

There are three types of active electromagnetic imaging systems for
data acquisition: the real-aperture (RA) system Fenn et al., 2000, the
synthetic-aperture (SA) system *Electronic Warfare and Radar Systems
Engineering Handbook* 1997; Charvat, 2014, and the random-pattern-
based imager Duarte et al., 2008; Watts et al., 2014, as illustrated in
Figure 4.2. The RA system composed of a great amount of elements in a
large aperture is flexible in controlling measurement modes, but suffers
from the big size, heavy weight, high power, and expensive price of
hardware. To overcome this drawback, several concepts of sparse array
were introduced, for instance, Bayesian sparse array Zhang et al., 2011c,
co-prime array Vaidyanathan and Pal, 2011; Tan et al., 2013, Cross-
Miller array Zhuge and Yarovoy, 2012, etc. Nonetheless, these sparse
arrays are still hampered from civil applications due to their impressive
cost. On the contrary, the SA system works with the mechanical move-
ment of a single sensor to form a virtual large-size scanning aperture
via data post-processing. The SA system has been playing a key role in
society of remote sensing, especially in the regimes of microwave and
millimeter wave. Apparently, the SA system has relatively low hardware cost compared to the RA system; however, it is inefficient in data acquisition besides the relatively low signal-to-noise ratio (SNR) due to the mechanical movement of the single sensor. Invoked by the CS theory, the concept of fixed single-sensor imaging system has been invented to promote the imaging performance in optics Duarte et al., 2008, terahertz Watts et al., 2014, and millimeter and microwave frequencies Watts et al., 2014; Yun et al., 2015. The operational principle common to these imaging systems is to encode the information of probed objects into the single detector through a sequence of spatial wavefield modulators with quasi-random patterns, followed by retrieving the information of probed objects through a sparse-model-based reconstruction algorithm. These systems sense the data in a compressed form rather than in Nyquist-Shannon manner. Therefore, they enable a large reduction potentially in the number of measurements for sensing the object that admits certain low-dimensional model. In this sense, we would like to refer to this type of imaging systems as the compressive imaging system. Note that a sequence of random modulators are manipulated in the sequential manner, leading to the data acquisition with relatively low efficiency. The single-sensor and single-shot imager by using the dispersive metamaterial has been reported by Hunt et al. for high resolution imaging Hunt and Smith, 2013 and by Li et al for high/super-resolution imaging Li et al., 2015, where a metamaterial-based lens is utilized to embody the spatial knowledge of imaged scene into one-dimensional temporal measurements. These systems pose the distinct advantage that the data acquisition could be real-time completed, providing the promising ability of tacking high-speed moving targets. Here, we would like to point out that the alliance of well-established programmable (or, reconfigurable) electromagnetic metamaterials with model-based signal processing opens a new exciting door for designing smarter imaging systems with lower hardware cost. The programmable metamaterials consist of an array of controllable units, where each metamaterial element is connected to an external control circuit such that its resonance can be damped by application of a bias voltage. Through applying different voltages to the control circuit, selected subsets of the elements
4.3 Applications of Model-based Electromagnetic Imaging

![Diagram](image)

**Figure 4.2:** (a) a high-directivity beam is raster scanned over some angular region and the reflected power is measured for each beam; many sources, each with a dynamic phase shifter and possibly an amplifier, comprise the large aperture. (b) SAR, a single transceiver element illuminates the target with a broad beam set by the aperture of the individual transceiver; by mechanically scanning the transceiver to many positions over an area, an aperture can be synthesized to achieve high resolution. In (c), a metamaterial aperture is used to produce a complex—e.g., pseudorandom—beam pattern that illuminates large portions of the target; the beam patterns change with random pattern, allowing scene spatial data to be encoded into single-sensor measurements Hunt and Smith, 2013

Spatial information of an imaging domain can be encoded onto this set of radiation patterns, or measurements, which can be processed to reconstruct the targets in the scene using low-dimensional-model based reconstruction algorithms. We would like to refer to Oliveri *et al.*, 2015 for comprehensive review on the topic of programmable metamaterials. Apparently, this kind of system has the advantage of simplicity and extremely low hardware cost compared to phased-array technique.

4.3 Applications of Model-based Electromagnetic Imaging

4.3.1 Basic SAR Imaging

SAR is a widespread active sensing technique in the area of computational imaging, which has the ability of all-time and all-weather observation. Conventional SAR system follows the guideline of Nyquist-Shannon theory, both in the data acquisition and in data processing.
Table 4.1: Parameters used in Figure 4.3

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center frequency</td>
<td>5.6GHz</td>
</tr>
<tr>
<td>Height of platform</td>
<td>3900m</td>
</tr>
<tr>
<td>Velocity of platform</td>
<td>About 108m/s</td>
</tr>
<tr>
<td>Antenna length</td>
<td>0.9m</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>80MHz</td>
</tr>
<tr>
<td>Sampling Rate</td>
<td>120MHz</td>
</tr>
<tr>
<td>Pulse duration</td>
<td>38us</td>
</tr>
<tr>
<td>PRF</td>
<td>Average 768Hz</td>
</tr>
</tbody>
</table>

The pulse repetition frequency (PRF) of SAR should be high enough to meet with the requirement of Nyquist-Shannon theory. In this way, the detectable swath scope of conventional SAR is constrained by the PRF, specifically, the higher the PRF is, the narrower the swath scope is Samadi et al., 2011. Another major drawback of conventional SAR is the low image resolution due to the use of back-propagation imaging algorithms. To address these issues, Cetin et al proposed the sparse-model-based SAR imaging method Cetin and Karl, 2001, which later on was generalized for wide-angle SAR imaging Varshney et al., 2006 and for auto-focusing SAR imaging Potter et al., 2010; Samadi et al., 2011. These results revealed that the sparse model, if correctly exploited, supports the enhanced SAR imaging performance, in particular, to get the enhanced imaging resolution from limited measurements Potter et al., 2010; Cetin and Karl, 2001; Samadi et al., 2011; Varshney et al., 2006. More importantly, the sparse model of underlying scene is helpful in expanding the scope of swath by reducing data sampling along the azimuth direction (i.e., lower PRF) Bingchen et al., 2012. In a word, the sparse-model-based SAR imaging provides the promising development of conventional SAR systems, which enables us to reduce the system complexity, save hardware cost, relieve the pressure from data transmission and storage, and improve the swath effectively.

Figure 4.3 demonstrates a set of airborne-SAR experiments, conducted in Sept. 2013, Tianjin, China, with different PRF, where ex-
4.3. Applications of Model-based Electromagnetic Imaging

Figure 4.3: Sparse-model-based SAR imaging results under different down-sampling PRF rate: (c) 50%; (d) 60%; (e) 70%; (f) 80%; (g) 90%; (h) 100%. This airborne-SAR experiment was conducted in Tianjin, China, in 2013, and the associated system parameters are summarized in Table 4.1. Additionally, the optical image is given in (a) and (b).
per experimental parameters are shown in Table 4.1. Note that the $l_0$-norm sparsity of undergoing scene is 9.5% in the wavelet transformed domain. It can be found from this set of figures that even the PRF is as low as 70% Nyquist rate, the imaging quality remains acceptable visually. Figure 4.4 provides the diagram of Donoho-Tanner phase transition for this used SAR configuration and associated sparsity-driven imaging algorithm, which is obtained using the Monte-Carlo simulation method with the addition noise level being SNR=30dB. It can be verified from Figure 4.4 that the sparse-model-based SAR imaging is capable of producing the image with the quality comparable to that using conventional SAR systems. Compared to conventional SAR imaging, the sparse-model-based SAR has superiority in noise and clutter reduction as well. Figure 4.5a and b compare two images obtained by the sparse-model-based SAR imaging and conventional SAR for detecting ships over sea, respectively. It can be observed that the sparse-model-based SAR imaging is superior to the conventional SAR imaging along with back-propagation algorithm, in terms of suppressing the oceanic clutter and noise. More specifically, the target-to-background ratios (TBRs) of Target1 and Target2 produced by sparsity-driven imaging are 41.92 dB and 48.34 dB, respectively, while these two values are 33.86 and 29.51 dB for conventional SAR imaging.

### 4.3.2 Tomographic SAR

The single-pass synthetic aperture radar (SAR) is a two-dimensional (2D) imaging technique, which is limited to produce the image of 3D objects in a 2D projection plan, i.e., the slant-range-Doppler plane. Tomographic SAR is a relatively new technology for 3D microwave remote imaging, which expands the traditional 2-D SAR imaging to the 3-D imaging by using multiple passes along the elevation dimension, as illustrated in Figure 4.6. Tomographic SAR is aimed at reconstructing 3D point clouds of discrete scatterers from only a few number of flight passes (usually in few tens) spaceborne/airborne SAR. An open question of tomographic SAR imaging is to solve the grating-lobe ambiguity along the elevation direction due to a few flight tracks available. Fortunately, the scatterers treated by tomographic SAR could be sparsely distributed
4.3. Applications of Model-based Electromagnetic Imaging

Figure 4.4: The phase-transition diagram corresponding to the SAR configuration and imaging algorithm used for Figure 4.3, which is calculated using the Monte-Carlo simulation over 500 random trials. The six white-filled circles from top to down in order mark the successful probabilities for the measurement with under-sampling rates of 100%, 90%, 80%, 70%, 60%, and 50%, respectively. Here, the additive SNR is 30dB. The blue region corresponds to the reliable solution while the red area for failed reconstruction.

Figure 4.5: Reconstructed images of oceanic ship targets using (a) the sparse-model-based SAR imaging and (b) conventional SAR imaging. The target-to-background ratios (TBRs) of Target1 and Target2 produced by sparsity-driven imaging are 41.92 dB and 48.34 dB, respectively, while these two values are 33.86 and 29.51 dB for conventional SAR imaging.
along the elevation dimension. This assumption is usually true for smooth 3D surface such as earth surface, even for complicated urban scenario Zhu and Bamler, 2014. Thus, the scene imaged by Tomo-SAR admits the sparse model. There are many encouraging results, which demonstrate that the sparse model of the scene under consideration supports the ability of Tomo-SAR imaging with super-resolution in the elevation dimension Zhu and Bamler, 2010.

Using simulated data, we show an example of using few repeat-pass tracks of SAR images to reconstruct 3D point clouds of a stadium. As shown in Figure 4.6, when the radar flies once along the x-direction over the scene of interest, each 2-D SAR image obtained is the projection of the scene over the azimuth-slant range plane (incident plane). If the radar makes multi-passes over the scene, that is, $N$ flights, a new synthetic aperture is formed along the s-direction (referred to as the elevation), which is normal to the $(x-r)$ plane. Then, the 3-D reflectivity function $\gamma(x, r, s)$ of the scene can be reconstructed from the multi-pass SAR data. After obtaining the 3-D scattering data, the 2-D azimuth-slant range data are first processed by the convolution back-projection imaging algorithm to get 11 frames of 2-D SAR images. These 11-pass images look almost the same, and one of them is shown in Figure 4.7. As shown in Figure 4.7, the 3-D stadium model has been projected onto the azimuth-slant range plane. As the radar illuminates the scene with an incident angle, strong scatterers are detected on the near-radar side of the model and weak scatterers on the off-radar side. The positions with higher scattering values in Figure 4.7 must have been overlapped with several scatterers, which need to be separated in the elevation direction. The final sparse-model-based 3-D reconstruction is shown in Figure 4.8, where the normalized scattering intensity above –50 dB is drawn in the figure. In this figure, the results have been converted from the imaging space to the target space, and the actual geometry model is shown for comparison.

4.3.3 ISAR Imaging

ISAR is a dual variant of SAR. It is powerful tool in imaging moving targets such as ships, aircrafts, and space objects by exploiting the
4.3. Applications of Model-based Electromagnetic Imaging

Figure 4.6: Geometry of Tomo-SAR imaging.

Figure 4.7: 2-D SAR image of the Beijing National Stadium model.
relative movement of target rather than radar motion to form a large observation aperture virtually. The high azimuth-resolution of ISAR imaging is constrained by the coherent processing interval (CPI) which reflects the virtual aperture length of relative cooperative movement between the radar platform and probed targets: the longer CPI is, the higher the resolution is. However, the long data acquisition may be unachievable in practice, since the electromagnetic scattering response of moving targets varies with the targets’ motion state, especially for the maneuvering or uncooperative targets. Therefore, a crucial issue to ISAR imaging is to achieve a high-resolution image from limited pulses (i.e., measurements). Fortunately, most of scenes treated by inverse SAR admits the use of sparse model since they consist usually of a few dominant point-like scatterers (i.e., scattering centers). As argued previously, the low-dimensional-model-based signal processing could sever as a valuable tool for ISAR imaging. On the other hand, ISAR achieves the high range resolution by using an ultra-band transmitting signal, which requires the high-performance A/D chip with very high
4.3. Applications of Model-based Electromagnetic Imaging

Figure 4.9: The ISAR imaging scheme of 3D space object in multi-orbit and multi-station modes.

sampling rate, if following the Nyquisit-Shannon theorem. In light of CS theory, the low-dimensional-model-based signal processing enables us to relieve the pressure on A/D sampling rate. Zhang et al studied the performance of the proof-of-concept CS based ISAR imaging, and demonstrated that the sparse-model-based signal processing supports the resolution improvement of ISAR imaging with limited data Zhang et al., 2009. In following years, the model-based ISAR imaging have been studied intensively for different application scenarios, for example, the 3D ISAR imaging using 2D sparse array Zhang et al., 2015, the stepped-frequency-waveform ISAR imaging Zhang et al., 2011b, the passive ISAR imaging Qiu et al., 1988, and so on. In this research branch, the study at its early stage is focused on the image scene that is sparse by itself. Recently, richer low-dimensional models have been studied, for instance, Wang et al Wang et al., 2014; Wang et al., 2015 and Duan et al Duan et al., 2015 explored the pattern-like model of the target scene to enhance the ISAR imaging quality. Table 4.2 summarizes some recent progress on this topic published in IEEE Xplore.

The examples below illustrate a simulated case of ISAR imaging of a 3D space object observed by a few ground stations in its multiple-
Table 4.2: Some progresses on structure-driven ISAR imaging

<table>
<thead>
<tr>
<th>Contribution</th>
<th>Algorithm</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application in Detecting micromotion targets via ISAR</td>
<td>Sparse Bayesian learning</td>
<td>Liu et al. (2014) Liu et al., 2014</td>
</tr>
<tr>
<td>Application in Monitoring maneuvering targets via ISAR</td>
<td>Smoothed SL0</td>
<td>Liu et al. (2013) Liu et al., 2013</td>
</tr>
<tr>
<td>Application in Bistatic ISAR</td>
<td>Iterative fixed-point method</td>
<td>Zhang et al. (2015) Zhang et al., 2015</td>
</tr>
<tr>
<td>Application in Fully polarimetric ISAR</td>
<td>L0-norm regularized method</td>
<td>Qiu et al. (2014) Zhang et al., 2015</td>
</tr>
<tr>
<td>Application in Passive bistatic ISAR</td>
<td>L0-norm regularized method</td>
<td>Qiu et al. (2015) Qiu et al., 1988</td>
</tr>
<tr>
<td>Application in autofocus ISAR</td>
<td>Iterative Newton method</td>
<td>Du et al. (2013) Du et al., 2013</td>
</tr>
<tr>
<td>Exploiting ISAR targets’ structure via random Markov fields</td>
<td>Sparse variational Bayesian learning</td>
<td>Wang et al. (2015) Wang et al., 2015</td>
</tr>
</tbody>
</table>
Figure 4.10: 3-D ISAR imaging results of KH-12 satellite. Here, we compares the results of original loss-less ISAR image (100% data) vs. 30%, 50%, 70% and 90% sampled aperture data-formed ISAR images.
orbit overpasses. The reference scheme is depicted in Figure 4.9. It is interesting to notice that this example is just an opposite problem of tomographic SAR, in the sense that the former is observing satellite from ground, while the latter is observing ground from satellite. Figure 4.9 shows the KH-12 satellite CAD model used to generate the raw radar observation signals. For 2D ISAR imaging, we purposely discard portion of the azimuth observations to mimic the real situation. Then the standard OMP algorithm is used to retrieve full aperture observation data, which is subsequently used to form 2D ISAR images through standard ISAR imaging algorithm. Figure 4.10 compares the results of original loss-less ISAR image (100% data) vs. 30%, 50%, 70% and 90% sampled aperture data-formed ISAR images. It demonstrates the fact that the level of details (LOD) of target can be recovered gradually with given portion of observation data.

Finally, despite encouraging results of model-based ISAR imaging, it is worthwhile pointing out that ISAR imaging is typically has the strong requirement on real-time processing, however the model-based ISAR imaging requires the implementation of iterative reconstruction algorithms, and thus is computationally inefficient compared to classical Fourier-based ISAR imaging methods. Therefore, it is desirable to develop efficient algorithms of model-based ISAR imaging to meet with the requirement on real-time detection.

4.3.4 Penetration Radar

Penetration radar is a tool of nondestructively detecting the concealed object embedded in an opaque structure, which has found a variety of important civil and military applications, such as, landmine detection, security inspection, search and rescue missions, to name a few Amin, 2011. Two popular members in this big family are the ground-penetrating radar (GPR) and the through-the-wall radar (TWR). In order to achieve high imaging resolution, both long (real or synthetic) observation aperture and ultra-wideband illumination signal are required. This results in a large amount of space-time or space-frequency data, long data acquisition time, as well as harsh storage requirement. Reduction of data volume is appealing in penetration radar imaging
including GPR and TWR, as it could speed up considerably data acquisition and data processing, leading to the reduction in both hardware and computational costs. More specifically, the conventional imaging algorithm adopted in penetration radar is almost the same as that used in previous radar imaging, like ISAR, TomoSAR, etc., where a large amount of well-sampled data over the line of Nyquist-Shannon theory is involved. As argued previously, the low-dimensional-model-based signal processing is a natural candidate to overcome the difficulties in the conventional method in penetration radar. Gurbuz et al. proposed that compressive sensing based GPR imaging is applicable when the target scene embedded in known underground is sparse, i.e. there only exist a small number of point targets Gurbuz et al., 2009b; Gurbuz et al., 2009a. It was demonstrated that the use of sparseness in the GPR target scene could significantly reduce the number of measurements, accelerate data acquisition, reduce image clutter and enhance the imaging resolution. In principle, TWR imaging is quite similar to GPR imaging except that the electromagnetic wave travels through different background media. In Yoon and Amin, 2008, CS was applied to TWR imaging in order to obtain the enhanced imaging performance from reduced measurements, where the wall parameters are assumed a prior known. Here, it should be emphasized that different from above free-space radar applications, a unique difficulty arising in the penetration radar is that the targets are embedded in the complicated but unknown environment, which makes the implementation of sparsity-driven processing algorithms more challenging, especially for the detection of concealed weak scattering targets. To address this challenge, some efforts have been made for TWR imaging applications; please refer to Lagunas et al., 2013; Ahmad et al., 2014; Leigsnering et al., 2015 and references therein. Some recent progresses on sparsity-driven penetration radar imaging are summarized in Table 4.3.

In the following some numerical and experimental results are presented for GPR and TWR imaging with CS. In the first example, we present the MIMO GPR imaging with limited transmitting/receiving elements and limited frequencies. In the simulation 17 transmitters are equally spaced from -0.96 m to 0.96 m and 16 receivers are equally
Table 4.3: Some progresses on sparsity-driven penetration radar imaging.

<table>
<thead>
<tr>
<th>Contribution</th>
<th>Algorithm</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application in imaging sparse indoor stationary scene</td>
<td>Orthogonal matching pursuit</td>
<td>Ahmad et al. (2015) Ahmad et al., 2014</td>
</tr>
<tr>
<td>Application in stepped-frequency GPR imaging</td>
<td>Interior point methods</td>
<td>Gurbuz et al. (2009) Gurbuz et al., 2009a Suksmono et al. (2010) Suksmono et al., 2010</td>
</tr>
<tr>
<td>Application in polarimetric TWI imaging</td>
<td>Sparse Bayesian learning</td>
<td>Wu et al. (2015) Wu et al., 2015</td>
</tr>
<tr>
<td>Application in MIMO array forward-looking GPR imaging</td>
<td>Basic pursuit declutter method</td>
<td>Yang et al. (2014) Yang et al., 2014</td>
</tr>
<tr>
<td>Application in topographic GPR imaging</td>
<td>The spectral projected gradient</td>
<td>Soldovieri et al. (2011) Soldovieri et al., 2011</td>
</tr>
<tr>
<td>Application in GPR blind deconvolution</td>
<td>Iterative reweighting method</td>
<td>Li (2014) Amin, 2011</td>
</tr>
<tr>
<td>Applications in imaging objects in complicated indoor environment</td>
<td>OMP</td>
<td>Leigsnering et al. (2015) Bingchen et al., 2012</td>
</tr>
<tr>
<td>Equivalent-source model-based iterative sparse imaging</td>
<td>L1-regularized CVX</td>
<td>Nikolic et al. (2012) Nikolic et al., 2012</td>
</tr>
<tr>
<td>Application in breast cancer localization</td>
<td>L1-regularized</td>
<td>Nikolic et al. (2016) Nikolic et al., 2010; Stevanovic et al., 2016b</td>
</tr>
</tbody>
</table>
spaced from -0.9 m to 0.9 m parallel to the ground at a height of 0.2 m. The dielectric constant, conductivity of the ground are $\varepsilon_b = 6$, $\sigma_b = 0.01 \text{S/m}$. The operating frequency covers the range from 0.8 GHz to 2 GHz with space of 24 MHz, i.e., the total number of discrete frequency is $N_f = 49$. The targets under investigation are a metallic cylinder and two rectangular metallic cylinders as shown in Figure 4.11 (a). Figure 4.11 (b) shows the imaging result using full data, i.e. data measured at all frequencies and all antenna locations. The true region of the target is indicated with white dashed circles. From this figure, we can clearly identify all the targets at their correct locations. To reduce the number of transmitters and receivers we selected randomly six transmitters from the total 17 transmitters and 6 receivers from the total 16 receivers. To reduce the number of measurement further, we randomly measure five frequencies at each receiver location. Figure 4.11 (c) shows the beamforming result using the limited measurement. Due to the lack of sufficient information about the target, the image is blurred, distorted, and has higher sidelobe levels, as evident from Figure 4.11 (c). By exploiting the sparsity of image scene and solving the sparse constraint optimization problem with the same limited data, we obtain the reconstructed image with CS as shown in Figure 4.11 (d). From the sparse-model-based MIMO GPR imaging result we find that the target can be clearly identified at its true position. From the comparison of Figure 4.11 (b) and (c) it is clear that a much higher resolution and cleaner image can be obtained through sparse constraint optimization.

Classical GPR images do not provide an accurate estimate of the target shape. The imaging problem is particularly challenging when the object of interest has a concave profile or sharp corners, which are the source of multiple scattering. In the next example, we focus on the imaging problem of complex-shaped targets hidden inside a dielectric body. The proposed approach defines a grid of equivalent sources (filament currents) uniformly spread in the dielectric interior Nikolic et al., 2013. The currents of the equivalent sources are unknown. However, we suppose that there are only a few equivalent sources with nonzero currents. Regularizing the $\ell_1$-norm emphasizes the equivalent sources in the vicinity of the target surfaces and thus provides information
about their shape. By exploiting the information provided by an array of antennas in a multiview/multistatic configuration, this method can reconstruct the target shapes using monochromatic scattered data. In the simulated example, a U-shaped perfectly conducting cylinder is embedded in a dielectric body with a square cross section and side length of 30 cm. The relative permittivity of the dielectric is $\varepsilon_r = 3 - j0.3$. The operating frequency is $f = 1$ GHz. The sensor array consists of 16 electric field probes, where the adjacent sensors are separated by 30 cm. The standoff distance of the array from the dielectric surface is 45 cm. Figure 4.12 shows the imaging results when the data is generated by a 2-D electromagnetic solver. The target contour is denoted by the black line. For comparison, we also applied the linear sampling method (LSM) Catapano et al., 2007 and MUSIC Agarwal and Chen, 2008. These two algorithms failed to recover the shape of the target as a consequence of the multiple scattering produced by the concave form of the cylinder. This simulations were validated via the following experiment. We consider a finite region limited by vertical aluminum
4.3. Applications of Model-based Electromagnetic Imaging

Figure 4.12: Imaging results of a U-shaped metallic cylinder hidden in a dielectric body computed from synthetic data (SNR = 20dB): (a) sparse imaging method; (b) linear sampling method (LSM); (c) MUSIC.

Figure 4.13: Imaging results of a U-shaped metallic cylinder hidden in a dielectric body computed from experimental data: (a) experimental setup; (b) sparse imaging method.

walls of 90 cm length. A dielectric cylinder with the embedded target is located at the center of the structure. The dielectric has a square cross section with a side length of 20 cm. The permittivity of the dielectric is \( \epsilon_r = 4 \) and the loss tangent is \( \tan \delta = 0.025 \). The sensors are thin wires with a diameter of 3 mm positioned as depicted in Figure 4.13 (a). The image that results from the experimental data can reconstruct the U shape of the target; however, it shows a few false targets due to the strong multipath generated by the metallic enclosure of the testbed.

In this example, a set of experimental TWR imaging results are presented. The measurement data are collected in an anechoic chamber as depicted in Figure 4.14. A pair of Archimedean spiral antennas is used to transmit and receive radar signals with the height of 1.3 m above the
Figure 4.14: Experiment scenario. A pair of Archimedean spiral antennas is used to transmit and receive radar signals with the height of 1.3 m above the ground. The stepped-frequency signal covers the 0.5-2.0 GHz frequency band with a step size of 1 MHz is used. The standoff distance from the wall is 1.38 m. The two trihedral reflectors are placed behind the wall with a mutual distance of 1.5 m on the ground. The fast spectral projection gradient $l_1$-norm method is used in CS reconstruction to yield the imaging result depicted in Figure 4.15, where the dashed circles indicate the true positions of the two trihedral reflectors. Low-order model-based imaging can offer finer resolution with only 6% data compared with back projection (BP) imaging result.

4.4 Future Trends

4.4.1 Single-sensor Computational Imager

Along the line of CS-driven imager in combination with programmable metamaterial, we propose a new compressive architecture here, as illustrated in Figure 4.16. Although this proposed imager is an instance of
4.4. Future Trends

Figure 4.15: Imaging results with the dynamic range of 10 dB: (a) BP imaging result with full data; (b) low-order model-based imaging result with 6% data.

the coded aperture imaging systems, it is different from the conventional CS-inspired imagers (e.g., the single-pixel camera and a recent terahertz single-sensor imager) where the elements of the random masks are manipulated in the pixel-wised manner. The controllable elements in this scheme are manipulated in the column-row-wised manner, which could greatly simplify the shutter control mechanism of the pixel-wised coded exposure. As illustrated in Figure 4.16, the proposed single-sensor imager is composed of three parts, including a transmitter working at single frequency to launch the incident wave, a 1-bit programmable coding metasurface to generate the sequentially CS random masks for modulating spatial wavefronts emanating from the transmitter, and a single sensor fixed at some distance away from the metasurface to collect the wave fields scattered from the probed object. The metasurface consists of a two-dimensional array of 1-bit voltage-controllable coding particles; and each particle could be in a state of two distinct responses: “1” when loaded with a high-level voltage for important secondary radiations, and “0” for almost negligible radiation. The coding elements in the proposed scheme are manipulated in the column-row-wised manner, rather than the pixel-wised manner in the conventional CS methods. Specifically, the coding metasurface with $N_x \times N_y$ particles is controlled by $N_x + N_y$ instead of $N_x \times N_y$ random binary sequences. In this way, the programmable coding metasurface is capable of producing quasi-random
Figure 4.16: The schematic of the proposed single-sensor imager: this imager consists of a transmitter working with single frequency launches an illumination wave, a one-bit coding metasurface is responsible for generating sequentially random masks for modulating the spatial wavefront emerged from transmitter, and a single sensor is fixed at somewhere for collecting the wave field scattered from the probed object.

patterns in a very flexible and dynamic manner.

The metasurface with the \( m \)th coded pattern, illuminated by an \( x \)-polarized plane wave, as illustrated in Figure 4.16, produces approximately an \( x \)-polarized radiation field at the location \( \mathbf{r} \):

\[
E^{(m)}(\mathbf{r}) = \sum_{n_x=1}^{N_x} \sum_{n_y=1}^{N_y} \tilde{A}^{(m)}_{n_x,n_y} g(\mathbf{r}, \mathbf{r}_{n_x,n_y}), \quad m = 1, 2, 3..., M \tag{4.4}
\]

where \( g(\mathbf{r}, \mathbf{r}_{n_x,n_y}) = \frac{\exp(jk_0|\mathbf{r}-\mathbf{r}_{n_x,n_y}|)}{4\pi|\mathbf{r}-\mathbf{r}_{n_x,n_y}|} \) is the three-dimensional Green’s function in free space, \( k_0 \) is the operation wavenumber, \( M \) is the total number of the coded patterns, \( \tilde{A}^{(m)}_{n_x,n_y} = A^{(m)}_{n_x,n_y} \exp(j\varphi^{(m)}_{n_x,n_y}) \) describes the induced \( x \)-polarized current with the amplitude \( A^{(m)}_{n_x,n_y} \) and phase \( \varphi^{(m)}_{n_x,n_y} \) on the \((n_x, n_y)\)th unit of the coding metasurface, and \( \mathbf{r}_{n_x,n_y} \) is the coordinate of the \((n_x, n_y)\)th metasurface particle. In Equation (4.4), the double summation is performed over all pixels of the coding metasurface, where \( n_x \) and \( n_y \) denote the running indices of the particle of the coding metasurface along the \( x \)- and \( y \)-directions, respectively. The
probed object with contrast function \( O(\mathbf{r}) \), falling into the investigation domain \( V \), is illuminated by the wave field of Eq. (4.4), giving rise to the following electrical field at \( \mathbf{r}_d \) Pastorino, 2010:

\[
E^{(m)}(\mathbf{r}_d) = \sum_{n_x=1}^{N_x} \sum_{n_y=1}^{N_y} \tilde{A}^{(m)}_{n_x,n_y} \int_{V} g(\mathbf{r}, \mathbf{r}_{n_x,n_y}) g(\mathbf{r}_d, \mathbf{r}) O(\mathbf{r}) \, d\mathbf{r} \quad (4.5)
\]

Note that the Born approximation has been implicitly used in Equation (4.5). After introducing the following function

\[
\tilde{O}_{n_x,n_y} = \int_{V} g(\mathbf{r}, \mathbf{r}_{n_x,n_y}) g(\mathbf{r}_d, \mathbf{r}) O(\mathbf{r}) \, d\mathbf{r}, \quad n_x = 1, 2, 3..., N_x, n_y = 1, 2, 3..., N_y \quad (4.6)
\]

Equation (4.5) becomes

\[
E^{(m)}(\mathbf{r}_d) = \sum_{n_x=1}^{N_x} \sum_{n_y=1}^{N_y} \tilde{A}^{(m)}_{n_x,n_y} \tilde{O}_{n_x,n_y}, \quad m = 1, 2, 3..., M \quad (4.7)
\]

By applying the far-field approximation, Eq. (4.6) can be rewritten as

\[
\tilde{O}_{n_x,n_y} = \left(\frac{1}{4\pi}\right)^2 \frac{e^{jk_0 r_{n_x,n_y}}}{r_{n_x,n_y}} \frac{e^{jk_0 r_d}}{r_d} \int_{V} e^{-j(k_{n_x,n_y} + k_d) \cdot \mathbf{r}} O(\mathbf{r}) \, d\mathbf{r} \quad (4.8)
\]

where \( k_{n_x,n_y} = k_0 \frac{r_{n_x,n_y}}{r_{n_x,n_y}} \) and \( k_d = k_0 \frac{r_d}{r_d} \). Eq. (4.8) implies that \( \tilde{O}_{n_x,n_y} \) corresponds to two-dimensional discrete Fourier transform of \( O(\mathbf{r}) \). Notice that the spatial bandwidth of \( O(\mathbf{r}) \) is limited by the maximum value of \(|k_{n_x,n_y} + k_d|\), and is determined by the maximum size of the coding metasurface. In other words, it can be deduced that the achievable resolution on \( O(\mathbf{r}) \) is in the order of \( O(\frac{\lambda R}{D}) \), where \( \lambda \) is the operating wavelength, \( R \) is the observation distance, and \( D \) is the maximum size of the coded aperture.

In the context of computational imaging, Equation (4.7) can be reformulated in the following compact form:

\[
E^{(m)} = \langle \tilde{\mathbf{A}}^{(m)}, \tilde{\mathbf{O}} \rangle \quad m = 1, 2, 3..., M \quad (4.9)
\]

where the symbol \( \langle \cdot \rangle \) denotes the matrix inner-product, the matrix \( \tilde{\mathbf{A}}^{(m)} \) has the size of \( N_x \times N_y \) with entries of \( \tilde{A}^{(m)}_{n_x,n_y} \), and the matrix \( \tilde{\mathbf{O}} \)
with the size of $N_x \times N_y$ is populated by $\tilde{O}_{n_x,n_y}$. Equation (4.9) reveals that the resulting problem of the computational imaging consists of retrieving $N = N_x \times N_y$ unknowns $\{\tilde{O}_{(n_x,n_y)}\}$ from the $M$ measurements $\{E^{(m)}(r_d)\}$. Typically, Equation (4.9) has no unique solution if $N > M$ due to its intrinsic ill-posedness. To overcome this difficulty, we pursue a sparsity-regularized solution to Equation (4.9) since we believe that the probed object $\tilde{O}$ has a low-dimensional representation in certain transform domain denoted by $\Psi$, i.e., $\Psi(\tilde{O})$ being sparse. Therefore, the solution to Equation (4.9) could be achieved by solving the following sparsity-regularized optimization problem:

$$
\min_{\tilde{O}} \left[ \frac{1}{2} \sum_{m=1}^{M} (E^{(m)} - \langle \tilde{A}^{(m)}, \tilde{O} \rangle)^2 + \gamma \|\Psi(\tilde{O})\|_1 \right]
$$

(4.10)

where $\gamma$ is a balancing factor to trade off the data fidelity and sparsity prior.

It is desirable that the number of sequential measurements of a fast microwave imaging should be as small as possible, and that the shutter control mechanism of the coded exposure should be as simple as possible. Traditionally, the coding pixels of the random mask are controlled independently, which is usually inefficient in data acquisition. Suffer from this fact, the temporal and spatial resolutions is limited. To break this bottleneck, a column-row-wised coding metasurface is proposed, by which the temporal random modulation of a $N_x \times N_y$ pixel array could be controlled by $N_x + N_y$ rather than $N_x \times N_y$ random binary sequences as required by a pixel-wise coding exposure. Specifically, the $N_x$-length random binary sequences of $\{0,1\}^{N_x}$ (denoted by $r = [r_1, r_2, \ldots, r_{N_x}]$) and the $N_y$-length random binary sequences of $\{0,1\}^{N_y}$ (denoted by $c = [c_1, c, \ldots, c_{N_y}]$) are used to control the row and column pixels, respectively. The row and column binary control signals jointly produce the binary random coded exposure sequence at the pixel location of $(n_x, n_y)$. Thus, one random realization of a coded pattern reads $\tilde{A} = r^T c$ up to a constant multiplicative factor. In this design, only $N_x + N_y$ control signals are needed to achieve the randomly coded exposures with $N_x \times N_y$ pixels, which could drastically reduce the complexity and increase the filling factor.
We now demonstrate that our single-sensor imager based on the 1-bit column-row-wised coding metasurface has a theoretical guarantee on successful recovery of a sparse or compressible object from its reduced measurements by solving a sparsity-regularized convex optimization problem. The conclusion is summarized in Theorem 1, and the proof of which is provided in Long-Gang Wang and Cui, 2017.

**Theorem 1:** With $M$, $N$ and $S$ defined as previously, a $S$-sparse $N$-length signal can be accurately retrieved with the probability not less than $2 \exp(-C(\log S \log N)^2)$, provided that the number of measurements with $M \geq C \cdot S \log \frac{N}{S}$ is up to a polynomial logarithm factor, where $C$ is a constant depending only on $S$.

We present a set of Monte-Carlo simulation results to verify Theorem 1 of the sampling theory. In this set of simulations, we consider the signals of dimension $N = N_x \times N_y = 32 \times 32$ whose nonzero entries are drawn i.i.d. from a standard normal distribution and are located on a support set drawn uniformly at random. We vary the number of random masks from $M = 8$ to $M = 1024$ with the step of size 8. Among all 100 trials run for each pair of $M$ and $S$, the ones that yield relative errors no more than $10^{-3}$ are counted as successful reconstruction. For comparison, the random mask programmed in the pixel-wised manner is investigated as well. Figure 4.17a and 4.17b show the phase transition diagrams of the estimation produced by Eq. (2) in terms of the number of masks $M$ and $S$ for the proposed column-row-wised and conventional pixel-wised coded masks, respectively. In these figures, the $x$- and $y$-axes correspond to the values of $M$ and $S$, respectively. It can be observed that the phase transition boundaries are almost linear in both models, agreeing with the relation between $M$ and $S$ suggested by Theorem 1. Although the phase transition boundary for the column-row-wised model is slightly lower (worse) than that for the pixel-wised model, the difference is not significant.

Regarding the fabrication of proposed 1-bit programmable metasurface in the microwave regime, it can be made is composed of $20 \times 20$ electric-LC (ELC) particles, shown in Figures 4.18a and 4.18b. Each
ELC particle has the size of $6 \times 6 \text{mm}^2$ printed on the top surface of the FR4-substrate with a dielectric constant of 4.3 and a thickness of 0.2mm, plotted in Figure 4.18b. The ELC particle is loaded with a pair of identical PIN biased diodes (SMP 1320-079LF), and the 1-bit states of “0” and “1” are controlled by the applied bias direct-current (DC) voltage: when the biased voltage is on a high level (3.3 V), this pair of diodes are ‘ON’; when there is no biased voltage, the diodes are ‘OFF’. The effective circuit models of the biased PIN diode at the ON and OFF states are illustrated in Figure 4.18c. To show this more clearly, we plot the transmission responses ($S_{21}$) of the ELC particle loaded with the voltage-controlled PIN diodes in Figure 4.18d, in which the effective circuit models of the biased PIN diode are incorporated into the CST Microwave Studio. This figure clearly shows that at the working frequency 8.3GHz, the ELC particle behaves as a ‘1’ element when the diodes are ON, and as a ‘0’ element when the diodes are OFF.

A set of proof-of-concept experiments have been conducted to verify the performance of the proposed single-sensor microwave imager. To accomplish this goal, we fabricate a sample of 1-bit programmable metasurface that is encoded in the column-row-wised manner, as shown in Figure 4.19a. In our imaging experiments, a vector network analyzer (VNA, Agilent E5071C) is used to acquire the response data by measuring the transmission coefficients ($S_{21}$). More specifically, a pair of horn antennas are connected to two ports of the VNA through two 4m-long
Figure 4.18: (a) The diagram of the 1-bit coding metasurface composed of $20 \times 20$ voltage-controlled ELC particles. For the visual purpose, the ELC particles coded with ON state are highlighted in blue, and the other are not highlighted. (b) The configuration for the voltage-controlled ELC unit, which has a period of 18 mm, and a size of $6 \times 6 \text{mm}^2$. The ELC unit is printed on a commercial printed circuit board FR4 with the relative permittivity of 4.3 and the thickness of 0.2 mm. (c) The effective circuit models of the biased diode at the ON and OFF states. (d) The $S_{21}$ responses of the ELC particle loaded with PIN diode, implying that the ELC particle behaves as a ‘1’ element when the diode is on, and as a ‘0’ element when the diode is off at the working frequency 8.3GHz.
50-Ω coaxial cables: one is used for launching the incident wave, and the other for collecting the response data emanated from the probed object, illustrated in Figure 4.19b. The experiments are carried out in a microwave anechoic chamber with the size of $2 \times 2 \times 2m^3$. Other relevant parameters are kept as the same as those adopted in the numerical simulations. These experiments are conducted to show the ability of the proposed coding metasurface to generate the controllable radiation patterns. The biased voltages of the coding metasurface in both column and row distributions could be digitally controlled by toggling different triggers, which control the ‘ON’ and ‘OFF’ states of the biased PIN diodes, thereby the required ‘0’ or ‘1’ state of each metasurface particle could be realized. Therefore, different quasi-random radiation patterns could be achieved. Totally 1000 random radiation patterns are generated for the imaging purpose, and three of them are shown in Figures 4.19d.

These radiation patterns are obtained by scanning the electrical fields at 3 mm away from the coding metasurface with a 50Ω-coaxial SMA tip, followed by the near-field-to-far-field transformation. The measured radiation patterns resemble those predicted by the numerical simulations, despite a little mismatches arising from the measurement errors, parasite effects of the diodes, and other possible reasons, which implies that the proposed 1-bit column-row-wise coding metasurface can be utilized to generate incoherent random-like masks. Second, two sets of imaging results are presented to demonstrate the performance of the proposed single-sensor imager. As done in the numerical simulations, the “P” and “K”-shaped metallic objects are considered in the experiments, shown in Figure 4.19c. The reconstruction results for “P” (“K”) are provided in Figure 4.19e-g (Figure 4.19h-j), considering different numbers of measurement $M = 200$, 400, and 600, respectively. Similar conclusions to the previous numerical simulations can be drawn immediately. The experimental imaging results clearly validate the feasibility of the proposed single-sensor imaging system based on the 1-bit column-row-wised programmable coding metasurface.

Here, we introduced a new single-sensor imager based on the 1-bit column-row-wised programmable coding metasurface for the high-rate-frame electromagnetic imaging. A sample of such a 1-bit coding
Figure 4.19: (a) Three coding patterns of the 1-bit column-row-wise encoded metasurface, in which the corresponding controlling signals for the column and row pixels are plotted in the left and top sides of each subfigure. Here, the yellow and green parts correspond to the “ON” and “OFF” elements, respectively. (b) Three radiation patterns of the metasurface corresponding to the coded patterns shown in (a). (c) The reconstruction results of the P-type metallic object for different measurement numbers $M = 200$, 400, and 600, from the left to the right. (d) The reconstruction results of the K-type metallic object for different measurement numbers $M = 200$, 400, and 600 from the left to right.
metasurface was fabricated and the proof-of-concept imaging tests were conducted in microwave frequencies to validate the single-sensor system for the high-frame-rate imaging, which could pave a new avenue for capturing and tracking moving objects, with either very high or very low speeds. The proposed single-sensor imager features two advantages over the existing methods: 1) avoiding the object dispersion due to the single-frequency reconstruction without the frequency agility; 2) encoding the programmable metasurface in the column-row-wise manner instead of the pixel-wise manner to reduce drastically the data acquisition time, resulting in the improved temporal and spatial resolutions. We also demonstrated that the very simple 1-bit column-row-wised coding metasurface has a theoretical guarantee to ensure that the required measurement number is comparable to that for the conventional pixel-wise encoded masks while maintaining nearly the same imaging quality. The proposed method is not only applicable for far-filed imaging, but also for near-field imaging, which will be studied in our future work. The new imaging system may be extended to the terahertz frequencies.

4.4.2 High-order Electromagnetic Imaging

As argued at the beginning of this chapter, high-resolution images result in ill-posed inverse problems. Compressive imaging techniques rely on the sparseness of the underlying signal of interest as the prior knowledge that helps regularizing the problem. In the former section, we presented a reconstruction algorithm, which uses a grid of equivalent currents. These currents are sparse if they are bound to a small portion of the region under test, where the target is located. Nevertheless, in real-life scenarios of electromagnetic imaging, the scattered field has significant spatial variations. Hence, it is hard to obtain an accurate image with a simple model. As a consequence of this information loss, targets with different shapes may have identical images.

In order to improve the imaging accuracy, we incorporate higher order basis functions into the sparse algorithm. In a two-dimensional (2D) space, an equivalent current is an infinitely long line current. Therefore, instead of line currents, we also consider current multipoles, i.e., dipoles, quadrupoles, etc. In contrast to the line source, multipoles
produce directive electric fields. Hence, by combining multipoles of different orders and orientations, we can obtain arbitrary scattered field patterns. The idea of using higher order sources has been also considered in Crocco et al., 2013, as a tool for enhancing the linear sampling method (LSM). It was demonstrated therein significant improvement in the shape estimation of complex targets, with respect to the standard formulation of LSM, by exploiting a suitable combination of the results achieved with the sources of different orders.

For the new procedure, we consider an arbitrary object illuminated by an incident electromagnetic field. For the sake of simplicity, we assume 2D geometry and transverse-magnetic (TM) polarization. In terms of the multipole expansion Stevanovic et al., 2016a, the electric field produced by the induced currents in the object is given by

\[
E(r) = \sum_{n=0}^{\infty} H_n^{(2)}(\beta r) [a_n \cos(n\phi) + b_n \sin(n\phi)]
\]

where \( H_n^{(2)} \) is the \( n \)th order Hankel function of the second kind, \( \beta \) is the wave number, \( a_n \) and \( b_n \) are the expansion coefficients, and \( r = (r, \phi) \) is the observation point in the polar coordinate system. In Eq. (4.11), the term associated with \( n = 0 \) corresponds to the field radiated by a line source, whereas the terms associated with \( n = 1 \) correspond to the fields radiated by two orthogonal line dipoles. Note that as the order \( n \) grows, the sine and cosine functions become more directive with respect to the variable \( \phi \). Assuming that the radius of the minimum circle covering the object is \( a \), only the components with indices smaller or equal to \( \tilde{n} = \beta a \) are significant. The components of the order of \( n > \tilde{n} \) have a low radiation efficiency, as they give rise to fields that vanish exponentially with the order.

From the spatial distribution of the equivalent sources or multipole sources with significant currents, we learn about the target shape. The locations of the equivalent sources, as well as their currents, are unknown. Therefore, as illustrated in Figure 4.20, we define as a search space an even grid of line currents of different orders. For each order \( n \), the electric field generated by the sources of the grid can be formulated in a linear, compact form as in Eq. (4.9), and then, the unknown currents
Figure 4.20: Grid of equivalent, multipole sources: (a) line currents \((n = 0)\); (b) two orthogonal dipoles \((n = 1)\).

Figure 4.21: Imaging results of a U-shaped metallic cylinder: (a) reconstruction using the standard sparse processing (only \(n = 0)\); (b) reconstruction using the multipoles of the orders \(n = 0, 8, 9\) at \(\text{SNR}=20\ \text{db}\); (c) idem at \(\text{SNR}=10\ \text{db}\); and (d) idem at \(\text{SNR}=5\ \text{db}\).

are estimated by solving the sparsity-regularized optimization problem given in Eq. (4.10). The final image results by superimposing the higher order images and the zero-order one.

As a simulated example of the approach, we consider a U-shaped cylindrical body, which is 26 cm wide and 34 cm tall. To compute the scattered electric field, we use the commercial software WIPL-D Pro assuming that the operating frequency is \(f=2\text{GHz}\). We consider 50 receiving antennas recording the scattered field. The simulated data are corrupted by additive Gaussian noise. For the inverse imaging problem, the search space is \(6\lambda \times 6\lambda\) and it is filled with a grid of 50 \(\times\) 50 nodes. Figure 4.21a shows the result of the sparse processing computed for \(\text{SNR}=20\ \text{dB}\). The zero-order multipoles estimate correctly the convex part of the body contour. As expected, the algorithm indicates strong scattering centers such as the wedges. However, irrespective of the value of the regularization parameter, the zero-order sources were not able to
reconstruct the concave part of the target. Moreover, due to the multiple scattering, a few false pixels appeared at the target opening. On the other hand, the use of the higher order models improves significantly the target image, as shown in Figure 4.21b-d for different noise levels. In particular, the erroneous pixels close to the opening of the U-shape were suppressed, thus allowing the appreciation of the concavity of the target. The simulated results are satisfactory even for a noise level as low as SNR=5 db.
In the past decade, we have witnessed that the low-dimensional models of imaged scenes, if exploited correctly, promote remarkably the performance of electromagnetic imaging. This survey provides a comprehensive overview on this emerging research branch, termed as the low-dimensional-model-based electromagnetic imaging (the model-based electromagnetic imaging, for short). Despite encouraging results in this area, several important issues relevant to practical applications remain open.

First, it is appealing to design the smarter system for real-time data acquisition, in which low complexity and low cost are desirable. Nowadays, most of imaging systems are designed to perform dedicated tasks, implying that a library of knowledge on imaged scenes can be readily obtained through simulations or/and experiments in advance. Invoked by this, the samples-aided data acquisition can be conceived to reduce significantly measurements, at the same time to enhance the imaging quality. Taking a millimeter-wave MIMO radar based human safety inspection as example, we can always get a library including a wide range of human body samples, by which we can design the array configuration for library-based radiation pattern, leading to more
efficient and adaptive data acquisition. Figure 5.1a shows a set of gray-scale human images out of total 11880 samples with different heights, at different locations, with/without carrying a threat weapon Liang et al., 2015. A PCA-guided antenna array has been designed to generate sampled-aided radiation pattern. Figure 5.1b compares the results with or without the use of the PCA-guided radiation pattern, where 200 total measurements are used, the leftmost one is the ground truth, the middle and rightmost correspond to with and without using PCA-guided patterns, respectively, which demonstrates the advantage of samples-aided data acquisition over conventional methods. Additionally, the on-line data processing is usually needed using the programmable chip like FPGA, DSP and other possible options assisted with controllable circuits mentioned previously.

Second, a fundamental question to be addressed is that how a low-dimensional model of imaged scene affects quantitatively the imaging performance. It is well accepted that the low-dimensional model of imaged scenes could improve the imaging quality; however, we still don’t know what the ultimate limit of such improvement is. Probably, the difficulty in answering this question is due to the gap between the low-dimensional model of imaged scene and the electromagnetic imaging, especially when the imaging resolution is relatively lower than variation scale of objects. More specifically, the physical mechanism between the electromagnetic wavefield and imaged scene is complicated nonlinear process in essence, however, the theories and methods of model-based signal processing developed so far are based on the linear model. In language of radar imaging, the radar scattering section of complicated object is usually a function of operational frequencies, view angle, polarization status, background medium, and so on. The concept of parametric or semiparametric scattering model may be a feasible strategy, for example, several scattering models based on canonical objects have been developed and widely used in the area of polarimetry SAR. Then, basic aspects of statistical theory are in place, including non-asymptotic bounds on the imaging resolution, the estimation error and the minimum measurements required for correct model selection.

Third, as we have frequently pointed out, the modern electromag-
Figure 5.1: (a) Five samples of human body with or without weapon out of 11880 PCA samples. (b) compares the results with or without the use of the PCA-guided radiation pattern, where 200 total measurements are used, the leftmost one is the ground truth, the middle and rightmost correspond to with and without using PCA-guided patterns, respectively. Liang et al., 2015
netic imaging involves massive measurements and high-dimensional variables. The efficient algorithms, like the first-order methods and their randomized variants, have not far reached the technical level for dealing with massive data and extracting finer structures involved in the very large-scale electromagnetic imaging, e.g., microwave remote sensing. More importantly, these algorithms, especially the randomized first-order strategies, have relatively low accuracy, leading directly to the degradation of imaging resolution and accuracy, and failing to capture finer details of interest. Even more, these low-accuracy algorithms will be responsible for image artifacts. Therefore, it becomes an urgent problem to develop novel imaging framework with the guarantee of higher image quality and lower computational efficiency. We made an important step toward this direction by converting the electromagnetic imaging problem into the problem of image processing via the physics-driven structured dimensional reduction and physics-driven patch-based image processing.

In conclusion, we believe that the model-based electromagnetic imaging is an active research direction, with particular emphasis to the design of smart imaging systems, efficient reconstruction algorithms and practical applications. This emerging research direction becomes much more relevant in the era of electromagnetic data deluge. It extends the reach of electromagnetic imaging theories and methods, and brings new relevance. This research direction will bridge the gap between the electromagnetic imaging and the low-dimensional-model-based signal processing, and will bring new relevance to the pioneering studies of the scientists and engineers that have laid the foundations of this exciting research area.
Acknowledgements

Li was funded by the National Natural Science Foundation of China (No. 61471006). Nehorai was funded by the AFOSR Grant No. FA9550-11-1-0210 and No. FA9550-16-1-0386.

The authors are grateful to Longgang Wang and Hengxin Ruan for their editorial work.


References


References


